J. Differential Equations 253 (2012) 2161–2190

Contents lists available at SciVerse ScienceDirect

Journal of Differential Equations

www.elsevier.com/locate/jde

Detection of standing pulses in periodic media by pulse interaction

Martina Chirilus-Bruckner [a], Guido Schneider [b,*]

[a] Department of Mathematics and Statistics, Boston University, 111 Cummington Street, Boston, MA 02215, USA
[b] Institut für Analysis, Dynamik und Modellierung, Universität Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart, Germany

ARTICLE INFO

Article history:
Received 31 December 2011
Available online 15 June 2012

MSC:
34C20
35L05
35Q60

Keywords:
Modulation equations
Bloch wave interaction
Photonics

ABSTRACT

Creation and manipulation of standing light pulses would pave the way for manufacturing all-optical computational devices. The most promising material type that is believed to support such standing pulses are periodically structured optical materials, so-called, photonic crystals. A way of detecting or even manipulating the position of a standing pulse is by the interaction with a second moving pulse. In this work we give a mathematical justification of the formulas which have been derived for interaction effects such as a shift of the pulse carrier and envelope. A significant part of the analysis is devoted to the proper definition of these quantities. The analysis is carried out for a nonlinear wave equation with spatially periodic coefficients which can be derived as a model for the description of the underlying physical system. Our method yields a separate description of internal and interaction dynamics and can even be employed for the analysis of the interaction of NLS-scaled wave packets in general dispersive systems with spatially homogeneous and spatially periodic coefficients.

Contents

* Corresponding author.
E-mail address: Guido.Schneider@mathematik.uni-stuttgart.de (G. Schneider).

0022-0396/$ – see front matter
http://dx.doi.org/10.1016/j.jde.2012.06.003

1. Introduction

Optical media consisting of a periodic arrangement of different dielectrics structured at a length scale comparable to the wavelength of light – so-called photonic crystals – are believed to support standing light pulses. Hence, they are promising materials for the realization of all-optical computational devices. A way of detecting or even manipulating the position of a standing pulse is by the interaction with a second moving pulse. In [14] formulas for interaction effects such as a shift of the pulse carrier and envelope have been derived. It is the purpose of this paper to give a mathematical justification of these formulas. This justification is carried out for a spatially periodic nonlinear wave equation

$$\partial_t^2 u = \partial_x^2 u - au + bu^3, \tag{1.1}$$

with $x, t, u(x,t) \in \mathbb{R}$ and where a, b are real-valued, smooth, 2π-periodic functions with $\inf_{x\in\mathbb{R}} a(x) > 0$. System (1.1) can be derived as a model for the description of the underlying physical system, see [9, Appendix B]. For the justification of the system which describes the interaction of the pulses we proceed as in [5] where we were able to separate internal and interaction dynamics of wave packets in nonlinear Schrödinger (NLS) scaling for the spatially homogeneous cubic Klein–Gordon equation

$$\partial_t^2 u = \partial_x^2 u - u + u^3 \tag{1.2}$$

up to high order. NLS scaling means that the wave packet obeys the scaling which allows us to derive the NLS equation for the description of slow modulations in time and space of the envelope of the underlying carrier wave.

However, there is a major difference in the analysis for (1.1) and (1.2). For (1.2) the derivation of the approximation equations for the description of the interaction dynamics can be made directly for the scalar equation. In contrast to that for (1.1) a deeper understanding of the situation is necessary in order to separate internal and interaction dynamics of the wave packets. Before we start with this more complicated case we recall what is known rigorously about pulse interaction in the spatially homogeneous case.

Terminology. Linearization is always meant around the trivial solution. We use the terms pulse and wave packets interchangeably. The wordings "at lowest order", "to high order", etc. are used without explicitly mentioning the small perturbation parameter $0 < \varepsilon \ll 1$. Constants which can be chosen independently of the small perturbation parameter very often are denoted with the same symbol C.

1.1. Pulse interaction in the spatially homogeneous case

For the cubic Klein–Gordon equation (1.2) we consider the initial profile given by

$$u(x,t)|_{t=0} = u_{pulse}(x,t)|_{t=0}, \quad \text{and} \quad \partial_t u(x,t)|_{t=0} = \partial_t u_{pulse}(x,t)|_{t=0}$$

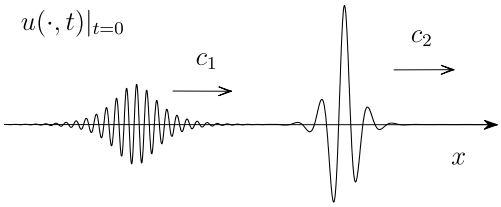

Fig. 1. Initial profile given by two wave packets traveling with different velocities.

with

$$u_{pulse}(x,t) = \varepsilon A_1\big(\varepsilon(x-c_1t),\varepsilon^2 t\big)e^{ik_1x+i\omega_1t} + \varepsilon A_2\big(\varepsilon(x-c_2t),\varepsilon^2 t\big)e^{ik_2x+i\omega_2t} + c.c., \tag{1.3}$$

where $c_j, k_j, \omega_j \in \mathbb{R}$, where $0 < \varepsilon \ll 1$ is a small perturbation parameter, and where the $A_j(\cdot,t)$ for $j = 1,2$ are localized functions for all $t \in \mathbb{R}$. See Fig. 1. The first observation is that single pulses really behave as predicted by the ansatz (1.3), cf. [8] for a rigorous approximation result. The same is true as long as the two pulses in NLS form stay apart. The second observation is that pulses with different velocities, i.e., $c_1 \neq c_2$, do not interact in lowest order w.r.t. the small perturbation parameter $0 < \varepsilon \ll 1$, i.e., in lowest order the first wave packet behaves as if the second wave packet was not present, and vice versa. This has been first rigorously proved in [11]. This property of non-interaction in lowest order has been discussed in [1] for various classes of original systems and general wave packets.

A different direction of research – and this is the one we persue in the present work – is concerned with the more detailed description of the interaction. In [10,14] formulas have been derived for the phase and envelope shift caused by the interaction of two pulses in soliton form. The first attempt to justify these formulas has been made in [4,13], where the shift of the envelope has been estimated to be of order $\mathcal{O}(\varepsilon)$ for pulses in soliton form. This had improved estimates known from [11,1] for the shift of the envelope from $\mathcal{O}(1)$ to $\mathcal{O}(\varepsilon)$.

In [5] the existing results have been improved significantly. For the Klein–Gordon equation (1.2) we were able to separate internal and interaction dynamics of general wave packets in NLS form up to high order. The internal dynamics is described by (uncoupled) nonlinear and linear Schrödinger equations. The interaction dynamics is mainly described by a phase shift function $\Omega_j^{(1)}$ of the underlying carrier wave and a pulse shift function $\Psi_j^{(1)}$ for the envelope.

In detail the interaction is approximately described by

$$\begin{aligned}
u(x,t) = {} & \varepsilon A_1^{(1)}\big(\varepsilon\big(x-c_1t+\varepsilon\Psi_1^{(1)}\big),\varepsilon^2 t\big)e^{i(k_1x+\omega_1t+\varepsilon\Omega_1^{(1)}+\varepsilon^2\Omega_1^{(2)})} \\
& + \varepsilon^2 A_1^{(2)}\big(\varepsilon\big(x-c_1t+\varepsilon\Psi_1^{(1)}\big),\varepsilon^2 t\big)e^{i(k_1x+\omega_1t+\varepsilon\Omega_1^{(1)}+\varepsilon^2\Omega_1^{(2)})} \\
& + \varepsilon^3 A_1^{(3)}\big(\varepsilon\big(x-c_1t+\varepsilon\Psi_1^{(1)}\big),\varepsilon^2 t\big)e^{i(k_1x+\omega_1t+\varepsilon\Omega_1^{(1)}+\varepsilon^2\Omega_1^{(2)})} \\
& + \varepsilon A_2^{(1)}\big(\varepsilon\big(x-c_2t+\varepsilon\Psi_2^{(1)}\big),\varepsilon^2 t\big)e^{i(k_2x+\omega_2t+\varepsilon\Omega_2^{(1)}+\varepsilon^2\Omega_2^{(2)})} \\
& + \varepsilon^2 A_2^{(2)}\big(\varepsilon\big(x-c_2t+\varepsilon\Psi_2^{(1)}\big),\varepsilon^2 t\big)e^{i(k_2x+\omega_2t+\varepsilon\Omega_2^{(1)}+\varepsilon^2\Omega_2^{(2)})} \\
& + \varepsilon^3 A_2^{(3)}\big(\varepsilon\big(x-c_2t+\varepsilon\Psi_2^{(1)}\big),\varepsilon^2 t\big)e^{i(k_2x+\omega_2t+\varepsilon\Omega_2^{(1)}+\varepsilon^2\Omega_2^{(2)})} \\
& + c.c.,
\end{aligned} \tag{1.4}$$

where $A_1^{(1)}$, $A_1^{(2)}$, $A_1^{(3)}$ describe the internal dynamics of the first wave packet and $A_2^{(1)}$, $A_2^{(2)}$, $A_2^{(3)}$ the internal dynamics of the second wave packet, i.e., $A_1^{(1)}$, $A_1^{(2)}$, $A_1^{(3)}$ behave as if the second wave packet

was not present and $A_2^{(1)}$, $A_2^{(2)}$, $A_2^{(3)}$ behave as if the first wave packet was not present. The interaction dynamics is described by the shift functions $\Psi_1^{(1)}$ and $\Psi_2^{(1)}$ for the envelopes and the shift functions $\Omega_1^{(1)}$, $\operatorname{Re}\Omega_1^{(2)}$, $\Omega_2^{(1)}$, and $\operatorname{Re}\Omega_2^{(2)}$ for the underlying carrier wave. $\operatorname{Im}\Omega_1^{(2)}$ and $\operatorname{Im}\Omega_2^{(2)}$ describe spatially localized modifications of the shape of the envelopes (it turned out that $\operatorname{Im}\Omega_j^{(1)} = 0$, i.e., there is no shape correction in first order). They decay rapidly to zero outside the region of interaction. The dependence of the interaction functions $\Psi_1^{(1)}, \dots, \Omega_2^{(2)}$ on the variables for space and time will be discussed in a moment. At this point we only remark that for $|x| \to \infty$ they converge to some constants.

This approximate description is accurate up to an error of order $\mathcal{O}(\varepsilon^{7/2})$, cf. [5, Theorem 4.2]. Hence, the effects (internal dynamics, envelope shift, phase shift) described by $A_1^{(1)}, \dots, A_2^{(3)}$ and $\Psi_1^{(1)}, \dots, \Omega_2^{(2)}$ can really be seen in the full system. A Taylor expansion of (1.4) w.r.t. ε shows that these effects are of order $\mathcal{O}(\varepsilon)$, $\mathcal{O}(\varepsilon^2)$, and $\mathcal{O}(\varepsilon^3)$ which are much larger than $\mathcal{O}(\varepsilon^{7/2})$ for $\varepsilon \to 0$, cf. Section 7.

Plugging in the ansatz (1.4) into the Klein–Gordon equation (1.2) and equating the coefficients in front of $\varepsilon^m e^{i(k_j x + \omega_j t + \varepsilon\Omega_j^{(1)} + \varepsilon\Omega_j^{(2)})}$ to zero gives a set of equations for the internal and interaction dynamics of the wave packets.

Internal dynamics. The starting equation for the description of the internal dynamics of the first wave packet is given by the NLS equation

$$-2i\omega_1 \partial_2 A_1^{(1)} = \big(1 - c_1^2\big)\partial_1^2 A_1^{(1)} + 3\big|A_1^{(1)}\big|^2 A_1^{(1)}.$$

The equations for the internal dynamics are completed by linear inhomogeneous Schrödinger equations for $A_1^{(2)}$ and $A_1^{(3)}$ which describe higher order corrections to the internal dynamics. For $r \in \{1, 2, 3\}$, the equations for the $A_1^{(r)}$ are completely uncoupled from the equations for the $A_2^{(r)}$ and from the variables that describe the interaction dynamics. The same is true vice versa for the equations for the $A_2^{(r)}$.

Interaction dynamics. The equations for the description of the interaction dynamics on the first wave packet start with the phase shift formula

$$\partial_1 \Omega_1^{(1)} = \frac{3}{k_1 - \omega_1 c_2}\big|A_2^{(1)}\big|^2,$$

followed by the envelope shift formula

$$\partial_1 \Psi_1^{(1)} = \frac{3(1 - c_1 c_2)}{(k_1 - c_2\omega_1)^2}\big|A_2^{(1)}\big|^2.$$

The equations are completed with a formula for $\Omega_j^{(2)}$ which describes the second order correction of the phase shift and amplitude which again can be solved by pure integration.

For $j = 1, 2$ the variables depend in the following way on the coordinates

$$\begin{aligned} A_j^{(m)} &= A_j^{(m)}\big(\varepsilon\big(x - c_j t + \varepsilon\Psi_j^{(1)}\big), \varepsilon^2 t\big), \\ \Omega_j^{(m)} &= \Omega_j^{(m)}\big(\varepsilon\big(x - c_{3-j} t + \varepsilon\Psi_{3-j}^{(1)}\big), \varepsilon^2 t\big), \\ \Psi_j^{(1)} &= \Psi_j^{(1)}\big(\varepsilon(x - c_{3-j} t), \varepsilon^2 t\big). \end{aligned}$$

In addition to the system of differential equations, we get a system of algebraic equations and corresponding non-resonance conditions from the cancelation of mixed and higher order harmonics which we refrain to discuss in the introduction.

The solutions to the equations which describe the internal and interaction dynamics exist globally in time for the following reasons. For the 1D NLS equation the global existence and uniqueness of solutions is well known due to its complete integrability. Given these solutions the other equations can be integrated. More precisely, the right hand sides in the equations for the $A_j^{(2)}$ (resp. $A_j^{(3)}$) depend at most linearly on $A_j^{(2)}$ (resp. $A_j^{(3)}$) such that global existence of solutions is guaranteed for them as well. The equations for $\Omega_j^{(1)}$ can be solved by pure integration. The same is then true for the governing equations for the higher order variables $\Psi_j^{(1)}$ and $\Omega_j^{(2)}$. The so far not displayed equations for the higher order harmonics are linear algebraic equations which can be solved explicitly.

In summary, for given solutions $A_1^{(1)}$, $A_2^{(1)}$ of the NLS equations all other equations can be integrated recursively yielding a very detailed and accurate description of the interaction mechanism.

Remark 1.1. So far we were not able to go to the next order of approximation, i.e., to give an even more detailed description. In the next order the terms could not be sorted into one of the baskets 'internal dynamics' or 'interaction dynamics'. For a more detailed discussion about this subject see [5]. Since this question should not be approached for the most general situation we skip any further discussion of it in the following.

1.2. Pulse interaction in the spatially periodic case

It is the purpose of this paper to describe pulse interaction in the spatially periodic case with the same preciseness as carried out for the spatially homogeneous case in [5]. Our goal is to separate internal and interaction dynamics of the wave packets up to the same order of accuracy. As already explained this is motivated by the idea to detect a standing pulse in a photonic crystal by the interaction with a second moving pulse [14].

Due to the 2π-periodic coefficient a in (1.1) the solutions of the linearized problem

$$\partial_t^2 u = \partial_x^2 u - au \tag{1.5}$$

are given by time-dependent multiples of Bloch modes, namely by

$$u(x,t) = \breve{u}_n(l,x)e^{ilx}e^{i\omega_n(l)t} \tag{1.6}$$

with $\breve{u}_n(l,x) = \breve{u}_n(l,x+2\pi) \in \mathbb{C}$ for $l \in [-1/2, 1/2)$ and $n \in \mathbb{Z} \setminus \{0\}$. The countably many curves of eigenvalues $l \mapsto \omega_n(l)$ are ordered to satisfy $\omega_n(l) = -\omega_{-n}(l)$.

Some curves exhibit horizontal tangencies at various wave numbers if they are plotted as a function over the Bloch wave numbers l. A wave packet with such a basic wave number, in the following called l_0, will have vanishing group velocity $c = d\omega_n/dl|_{l=l_0} = 0$ and so, in principle, standing wave packets are possible. In lowest order these wave packets are given by

$$u_{pulse}(x,t) = \varepsilon A\big(\varepsilon(x-ct), \varepsilon^2 t\big)e^{il_0x + i\omega_n(l_0)t}\breve{u}_n(l_0,x) + \mathcal{O}\big(\varepsilon^2\big) + c.c. \tag{1.7}$$

The approximation result [3, Theorem 1.1] shows that the evolution of the wave packet is described in lowest order via the solutions of an NLS equation correctly up to an error $\mathcal{O}(\varepsilon^2)$ on a time scale of order $\mathcal{O}(1/\varepsilon^2)$. Hence, in principal photonic crystals can be used as optical storage where the standing wave packet stands for a one in digital encoding of information. In fact, if A is close to an NLS soliton the results from [9,2] guarantee the existence of the standing wave packet on much longer time scales. One idea to detect or even manipulate the position of a standing wave packet is by the interaction with a second moving one.

When transferring results from [5] to the spatially periodic case the first guess would be to keep the above ansatz, but modified with $\breve{u}_{n_j}(l_j,x)$, i.e., to take

$$\begin{aligned} u_{pulse}(x,t) = {} & \varepsilon A_1^{(1)}\big(\varepsilon\big(x - c_1 t + \varepsilon\Psi_1^{(1)}\big), \varepsilon^2 t\big) e^{il_1 x + i\omega_{n_1}(l_1)t + i\varepsilon\Omega_1^{(1)} + i\varepsilon^2\Omega_1^{(2)}} \breve{u}_{n_1}(l_1, x) + \cdots \\ & + \varepsilon A_2^{(1)}\big(\varepsilon\big(x - c_2 t + \varepsilon\Psi_2^{(1)}\big), \varepsilon^2 t\big) e^{il_2 x + i\omega_{n_2}(l_2)t + i\varepsilon\Omega_2^{(1)} + i\varepsilon^2\Omega_2^{(2)}} \breve{u}_{n_2}(l_2, x) + \cdots \\ & + c.c. \end{aligned}$$

However, expanding this ansatz w.r.t. ε shows that in order to be precise up to order $\mathcal{O}(\varepsilon^3)$ the change of the eigenfunctions $\breve{u}_{n_j}(l,x)$ w.r.t. l has to be taken into account, too. Using the expansions

$$\breve{u}_{n_j}(l,x) = \breve{u}_{n_j}(l_j,x) + \partial_l \breve{u}_{n_j}(l_j,x)(l-l_j) + \partial_l^2 \breve{u}_{n_j}(l_j,x)(l-l_j)^2/2 + \mathcal{O}(l-l_j)^3$$

and using that multiplication with wave numbers in Bloch space correspond in physical space to derivatives, a more appropriate ansatz is given by

$$\begin{aligned} u_{pulse}(x,t) = {} & \big(\breve{u}_{n_1}(l_1,x) + \varepsilon\partial_l\breve{u}_{n_1}(l_1,x)\partial_{X_1} + \varepsilon^2\partial_l^2\breve{u}_{n_1}(l_1,x)\partial_{X_1}^2/2\big) \\ & \times \big(\varepsilon A_1^{(1)}\big(\varepsilon\big(x - c_1 t + \varepsilon\Psi_1^{(1)}\big), \varepsilon^2 t\big) e^{il_1 x + i\omega_{n_1}(l_1)t + i\varepsilon\Omega_1^{(1)} + \mathcal{O}(\varepsilon^2)} + \cdots\big) \\ & + \big(\breve{u}_{n_2}(l_2,x) + \varepsilon\partial_l\breve{u}_{n_2}(l_2,x)\partial_{X_2} + \varepsilon^2\partial_l^2\breve{u}_{n_2}(l_2,x)\partial_{X_2}^2/2\big) \\ & \times \big(\varepsilon A_1^{(2)}\big(\varepsilon\big(x - c_2 t + \varepsilon\Psi_2^{(1)}\big), \varepsilon^2 t\big) e^{il_2 x + i\omega_{n_2}(l_2)t + i\varepsilon\Omega_2^{(1)} + \mathcal{O}(\varepsilon^2)} + \cdots\big) \\ & + c.c. \end{aligned}$$

This ansatz is at least hard to guess. Interestingly this problem of wave number dependent eigenfunctions not only occurs in the spatially periodic case. It even appears in spatially homogeneous dispersive wave systems like the Maxwell–Lorentz system or the water wave problem. In fact, the novelty of this paper is not (only) the now possible handling of the spatially periodic situation, but also the handling of the general case of wave number dependent eigenfunctions. Throughout the paper we will explain how the method which we present for (1.1) also allows to separate internal and interaction dynamics of wave packets in NLS scaling for general dispersive systems with spatially homogeneous and spatially periodic coefficients.

1.3. The plan of the paper

There are different ways to approach the separation of internal and interaction dynamics for the interaction of two wave packets. We choose to stay as close as possible to [5] where this problem has been addressed in the setting of a Klein–Gordon equation with constant coefficients. The approach we present here might at first sight not seem the most direct one, however, we strongly believe that it is the most simple and, in the end, also the most natural one. It allows for generalizations in the sense that it gives an algorithm to handle more general situations. To be more specific, our strategy is to first transform the system into what we like to call "natural coordinates" and then carry out the separation of internal and interaction dynamics as presented in [5] for a two-dimensional (or four-dimensional) subsystem of the transformed system. This procedure discloses the similarity of the pulse interaction process for various equations, which really came as a surprise and might not be evident from a conventional perturbation approach, i.e., without a prior transformation of a given system many more terms might be created when inserting the ansatz into the system and in most situations it is not obvious at all if they can be eliminated by a more elaborate ansatz or if they belong to the internal dynamics or to the interaction dynamics.

Therefore, the plan to tackle this problem and the plan of the paper is as follows. Section 2 describes all transformations and approximations of Eq. (1.1) that enable us to transfer the procedure from [5] to equations with periodic coefficients: In Section 2.1 we write (1.1) as in [3] as a diagonalized first order system in Bloch space. In Section 2.3 the infinite-dimensional first order system

is then cut off in Bloch space such that it can be interpreted as an infinite-dimensional first order system in Fourier space. It turns out that it is sufficient to consider a two- or four-dimensional first order subsystem which can be handled more or less exactly as the spatially homogeneous cubic Klein–Gordon equation (1.2). Only for the handling of the higher order harmonics the complete system has to be considered. In Section 2.2 we explain how general dispersive wave systems with spatially homogeneous and spatially periodic coefficients can be brought into the form of the system derived in Section 2.1. For the formal derivation of amplitude equations in Section 3 the system from Section 2 is then transformed back into x-space. The formal error, which is also called the residual, is estimated in Sections 4 and 5. In Section 6 we prove the validity of this approximation of pulse interaction based on the estimates of the previous sections combined with a simple application of Gronwall's inequality back in the original system. Finally, in Section 7 a discussion of the result w.r.t. applications is given.

2. Transformation to "natural coordinates"

An immediate transfer of the findings from [5] to the spatially periodic case turned out to be difficult in two ways: On the one hand we have seen that the dependence of the eigenfunctions w.r.t. to the Bloch wave numbers makes it hard to guess the correct ansatz in x-space (see Section 1.2). On the other hand the Fourier or Bloch transform of the ansatz (1.4) is rather intransparent. The idea to solve this dilemma is as follows. First we proceed as in [3] where (1.1) has been written as a diagonalized first order system in Bloch space (system (2.6)). Secondly, the infinite-dimensional first order system is cut off in Bloch space such that it can be interpreted as an infinite-dimensional first order system in Fourier space. The expressions associated to the latter system in x-space only involve constant coefficient pseudo differential operators and are therefore well suited for an ansatz according to (1.4) that yields a separate description of internal and interaction dynamics of two interacting wave packets. We will devote a subsection to each of the afore mentioned steps. Furthermore, we will also discuss in a separate subsection how one carries over the techniques presented here to general nonlinear dispersive systems with spatially homogeneous and spatially periodic coefficients.

2.1. Bloch wave transform and diagonalization

The Bloch transform

$$\mathcal{B}(u)(l,x) = \breve{u}(l,x) = \sum_{n\in\mathbb{Z}} \hat{u}(l-n)e^{inx}, \quad \hat{u} = \mathcal{F}(u),$$

is an adaption of Fourier analysis in the spatially homogeneous case to the spatially periodic case (cf. [16]). It satisfies

$$\breve{u}(l,x) = \breve{u}(l,x+2\pi) \quad \text{and} \quad \breve{u}(l+1,x) = \breve{u}(l,x)e^{ix},$$

and its inverse is given by

$$\mathcal{B}^{-1}(\breve{u})(x) = u(x) = \int_{-1/2}^{1/2} \breve{u}(l,x)e^{ilx}\,dl.$$

Following [3] we apply the Bloch transform to (1.1) which gives

$$\partial_t^2 \breve{u} = -\breve{L}\breve{u} + b\breve{u}^{\star 3},$$

where

$$(\check{L}\check{u})(l,x) = \check{L}(l,\partial_x)\check{u}(l,x) = -\big((\partial_x + il)^2 - a(x)\big)\check{u}(l,x),$$

and where the operation $\star$ is the Bloch convolution

$$(\check{u} \star \check{v})(l) = \int_{-1/2}^{1/2} \check{u}(l-m)\check{v}(m)\,dm,$$

whose $(n-1)$-times iteration is abbreviated by $\check{u}^{\star n} = \underbrace{\check{u} \star \cdots \star \check{u}}_{n \text{ times}}$.

For fixed $l \in [-\frac{1}{2}, \frac{1}{2})$ the operator $\check{L}(l,\partial_x)$ is self-adjoint and positive definite in the space $L^2((0,2\pi],\mathbb{C})$. Since for each fixed l its resolvent is compact there are countably many eigenvalues $\lambda_j(l) > 0$ with corresponding eigenfunctions $f_j(\cdot,l)$ which form a Schauder basis of $L^2((0,2\pi],\mathbb{C})$. Furthermore, the eigenfunctions $f_j(\cdot,l)$ can be chosen to depend smoothly on l (cf. [3, Section 3.2]). An expansion of $\check{u}(l,\cdot,t)$ w.r.t. the set of eigenfunctions possesses the following analytic property.

Definition 2.1. We define the space

$$\ell^2(s) = \big\{\tilde{u} : \mathbb{N} \to \mathbb{C} \;\big|\; \|\tilde{u}\|_{\ell^2(s)} < \infty\big\}$$

equipped with the norm

$$\|\tilde{u}\|_{\ell^2(s)} = \Big(\sum_{j\in\mathbb{N}} \big|(1+j^2)^{s/2}\,\tilde{u}_j\big|^2\Big)^{1/2}$$

where $\tilde{u} = (\tilde{u}_j)_{j\in\mathbb{N}}$.

Lemma 2.2. *Let $l \in [-1/2, 1/2)$ be fixed. The mapping*

$$\mathcal{P}_l : H^s\big((0,2\pi],\mathbb{C}\big) \to \ell^2(s), \quad v \mapsto \big(\langle v(\cdot), f_j(l,\cdot)\rangle\big)_{j\in\mathbb{N}},$$

with $\langle u, v\rangle = \int_0^{2\pi} \overline{u(x)}v(x)\,dx$, is an isomorphism for each $s \geqslant 0$. Moreover, the mapping

$$\mathcal{P} : L^2\big([-1/2,1/2), H^s\big((0,2\pi],\mathbb{C}\big)\big) \to L^2\big([-1/2,1/2), \ell^2(s)\big),$$

with $\mathcal{P} : [-1/2,1/2) \mapsto \mathcal{P}_l$, is an isomorphism for each $s \geqslant 0$.

Proof. See [3, Lemma 3.3]. □

Now for each fixed $l \in [-1/2, 1/2)$ we expand $\check{u}$ w.r.t. the family of eigenfunctions $(f_j(l,\cdot))_{j\in\mathbb{N}}$, i.e.

$$\check{u}(l,x,t) = \sum_{j\in\mathbb{N}} \tilde{u}_j(l,t)\,f_j(l,x) = \mathcal{P}^{-1}\big((\tilde{u}_j)_{j\in\mathbb{N}}\big)(l,x,t)$$

and

$$\big(\tilde{u}_j(l,t)\big)_{j\in\mathbb{N}} = \big(\langle \check{u}(l,\cdot,t), f_j(l,\cdot)\rangle\big)_{j\in\mathbb{N}} = \mathcal{P}(\check{u})(l,t).$$

The equations for the coefficient functions $\tilde{u}_j$ are then given by

$$\partial_t^2 \tilde{u}_j(l,t) = -\omega_j^2(l)\tilde{u}_j(l,t) + \big\langle b(\cdot)\big[\mathcal{P}^{-1}\big((\tilde{u}_j)_{j\in\mathbb{N}}\big)\big]^{\star 3}(l,\cdot,t),\, f_j(l,\cdot)\big\rangle,$$

with $\omega_j(l) = \sqrt{\lambda_j(l)}$ for $j \in \mathbb{N}$, $l \in [-1/2, 1/2)$. Following [3, Section 3.2] these infinitely many second order systems are written as first order systems by introducing the variables $\tilde{Z}_j = (\tilde{Z}_{j,1}, \tilde{Z}_{j,2})^T$ through

$$\tilde{Z}_{j,1}(l,t) = \tilde{u}_j(l,t) \quad \text{and} \quad \tilde{Z}_{j,2}(l,t) = \frac{1}{\omega_j(l)}\partial_t \tilde{u}_j(l,t).$$

Doing so we arrive at

$$\partial_t \tilde{Z}_j(l,t) = \tilde{A}_j(l)\tilde{Z}_j(l,t) + \hat{N}_j\big((\tilde{Z}_n)_{n\in\mathbb{N}}\big)(l,t)$$

where

$$\tilde{A}_j(l) = \begin{pmatrix} 0 & \omega_j(l) \\ -\omega_j(l) & 0 \end{pmatrix},$$

and

$$\hat{N}_j\big((\tilde{Z}_n)_{n\in\mathbb{N}}\big)(l,t) = \frac{1}{\omega_j(l)}\big\langle b(\cdot)\big(\mathcal{P}^{-1}\big((\tilde{Z}_{n,1})_{n\in\mathbb{N}}\big)\big)^{\star 3}(l,\cdot,t),\, f_j(l,\cdot)\big\rangle e_2,$$

with $e_2 = (0,1)^T$. Finally, the transformation $\tilde{V}_n = Q^*\tilde{Z}_n$, where

$$Q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix},$$

diagonalizes the linear part of our system and brings it to the form

$$\partial_t \tilde{V}_j(l,t) = \tilde{D}_j(l)\tilde{V}_j(l,t) + \tilde{N}_j\big((\tilde{V}_n)_{n\in\mathbb{N}}\big)(l,t), \tag{2.1}$$

with linear part

$$\tilde{D}_j(l) = \begin{pmatrix} i\omega_j(l) & 0 \\ 0 & -i\omega_j(l) \end{pmatrix},$$

and nonlinear part

$$\begin{aligned}\tilde{N}_j\big((\tilde{V}_n)_{n\in\mathbb{N}}\big)(l,t) &= \frac{1}{\omega_j(l)}\big\langle b(\cdot)\big(\mathcal{P}^{-1}\big(\big([Q\,\tilde{V}_n]_1\big)_{n\in\mathbb{N}}\big)\big)^{\star 3}(l,\cdot,t),\, f_j(l,\cdot)\big\rangle Q^* e_2 \\ &= \frac{i}{4\,\omega_j(l)}\big\langle b(\cdot)\Big(\sum_{n\in\mathbb{N}}\big(\tilde{v}_n(\cdot,t) + \tilde{v}_{-n}(\cdot,t)\big)f_n(\cdot,x)\Big)^{\star 3}(l,\cdot,t),\, f_j(l,\cdot)\big\rangle(-1,1)^T,\end{aligned}$$

where we used the notation

$$\tilde{V}_n := (\tilde{v}_n, \tilde{v}_{-n})^T$$

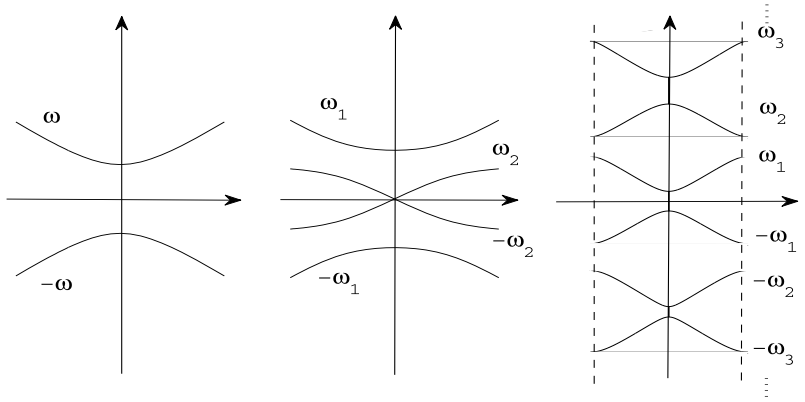

Fig. 2. *Left panel*: Dispersion relation of a cubic Klein–Gordon equation (1.2); *Middle panel*: Dispersion relation of a Maxwell–Lorentz system (2.2)–(2.3); *Right panel*: Dispersion relation of a cubic Klein–Gordon equation with periodic coefficients (1.1).

and dissolved $Q^* e_2 = (i/\sqrt{2})\,(-1,1)^T$ and $[Q\,\tilde{V}_n]_1 = (1/\sqrt{2})\,(\tilde{v}_n + \tilde{v}_{-n})$, where $[\cdot]_j$ is the j-th component of a vector. This system has now a diagonal linear part and a nonlinearity in convolution form. This is what we like to refer to as "system in natural coordinates". Hence, we are done with the first step of our program.

2.2. Longer remark about the general case

By the subsequent sections it will be clear that whenever we are able to transform a system into "natural coordinates", i.e., into a system with a diagonal linear part and a nonlinearity in convolution form, at least formally the internal and interaction dynamics can be separated. As already mentioned the problem of wave number dependent eigenfunctions not only occurs in the spatially periodic case. It appears in general spatially homogeneous dispersive wave systems, too. Thus, the novelty of this paper is not (only) the now possible handling of the spatially periodic situation, but the now possible handling of the general case of wave number dependent eigenfunctions. The reader which is mainly interested in the handling of (1.1) may skip this section in a first reading and proceed directly with Section 2.3.

As instructive examples we consider three original systems, namely the spatially homogeneous Klein–Gordon equation (1.2) which has been handled in [5], the spatially periodic Klein–Gordon equation (1.1) which has been transformed in the previous sections and the Maxwell–Lorentz system as an example for a homogeneous dispersive wave system.

The Maxwell–Lorentz system

$$\partial_t^2 u = \partial_x^2 u + \partial_t^2 p, \tag{2.2}$$

$$\partial_t^2 p = -\omega_0^2 p - d^2 u - p^3, \tag{2.3}$$

with $x, t, u = u(x,t),\ p = p(x,t) \in \mathbb{R}$ and $\omega_0, d, \gamma \in \mathbb{R}$ some constants describes the evolution of the electric field u and of the polarization p in some dielectric medium. Part (2.2) comes from Maxwell's equations while part (2.3) is a constitutive law in which for every $x \in \mathbb{R}$ the polarization satisfies equations for a driven nonlinear oscillator. The dispersion relation of the Maxwell–Lorentz system is given by

$$k^2 = \omega^2 + d^2 \frac{\omega^2}{\omega_0^2 - \omega^2}.$$

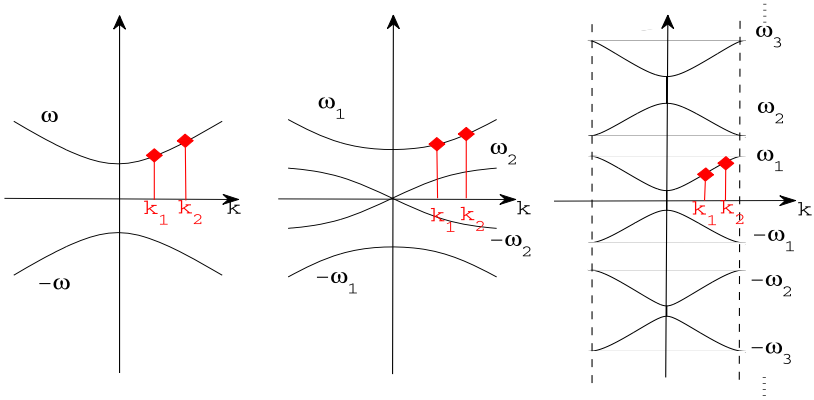

Fig. 3. Concentration of spectral content for two-pulse solutions.

See middle panel of Fig. 2. In [5] we discovered how to carry out the pulse interaction analysis for the cubic Klein–Gordon equation (1.2) whose dispersion relation has just one pair of branches of curves of eigenvalues (see left panel of Fig. 2). The following observation paves the way to transferring the findings to a more general setting. We will extensively make use of this property in Section 2.3.

Concentration of the spectral content of wave packets. Consider the Fourier transform of the ansatz for a single pulse

$$\begin{aligned}
&\mathcal{F}\{\varepsilon A_1(\varepsilon(\cdot+\omega_1'(k_1)t),T)e^{ik_1\cdot+i\omega_1(k_1)t}\}(k)\\
&\quad=\varepsilon\int_{\mathbb{R}} A_1(X_1,T)e^{ik_1x+i\omega_1(k_1)t}e^{-ikx}\,dx\\
&\quad=\varepsilon\int_{\mathbb{R}} A_1(X_1,T)e^{i\frac{k_1-k}{\varepsilon}\varepsilon x}\,dx\,e^{i\omega_1(k_1)t}\\
&\quad=\varepsilon\varepsilon^{-1}\int_{\mathbb{R}} A_1(X_1,T)e^{-iK_1X_1}\,dX_1\,e^{i\omega_1(k_1)t-i\varepsilon K_1\omega'(k_1)t}\\
&\quad=\hat{A}_1(K_1)\,\hat{E}_1
\end{aligned}\tag{2.4}$$

where we used the notation $X_1=\varepsilon(x+\omega_1'(k_1)t)$, $K_1=\frac{k-k_1}{\varepsilon}$, $\hat{E}_1=e^{i\omega_1(k_1)t-i\varepsilon K_1\omega'(k_1)t}$.

Remark 2.3. Note that when we switch from x-space to Fourier space, a factor ε^{-1} occurs due to the integration involved in the Fourier transform.

As a consequence, due to the concentration in Fourier or Bloch space a two-pulse ansatz for a more complicated system will in fact also only "feel" the pair(s) of branches to which the temporal frequencies $\omega_{n_j}(k_j)$ belong (see Fig. 3). Intuitively, as we already explained, regardless of the complexity of the dispersion relation, we can always restrict the analysis either, if $n_1=n_2$, to a two-dimensional subsystem, or, if $n_1\neq n_2$, to a four-dimensional subsystem belonging to the respective pair(s) of branches. The other branches only have to be considered for the computation of the higher order harmonics. Such a reduction would directly enable us to carry over the methods from [5]. However, in order to implement this reasoning to a given equation, we will have to perform a transformation

that brings it to a form where the branches of the dispersion relation are explicitly visible. It is this form that we referred to as "system in natural coordinates". It has the general structure

$$\begin{aligned}\partial_t \hat{V}_1 &= \begin{pmatrix} i\omega_1 & 0 \\ 0 & -i\omega_1 \end{pmatrix} \hat{V}_1 + \hat{N}_1\big[(\hat{V}_j)_{j\in\mathcal{N}}\big],\\ \partial_t \hat{V}_2 &= \begin{pmatrix} i\omega_2 & 0 \\ 0 & -i\omega_2 \end{pmatrix} \hat{V}_2 + \hat{N}_2\big[(\hat{V}_j)_{j\in\mathcal{N}}\big],\\ &\vdots\end{aligned}$$

or, written in a more compact form,

$$\partial_t \hat{V}_n = \hat{D}_n \hat{V}_n + \hat{N}_n\big[(\hat{V}_j)_{j\in\mathcal{N}}\big], \tag{2.5}$$

for $\hat{V}_n := (\hat{v}_{n,1}, \hat{v}_{n,2})^T$, $\hat{v}_{n,l} = \hat{v}_{n,l}(k,t) \in \mathbb{C}$, $n \in \mathcal{N}$, $l = 1,2$, where $\mathcal{N}$ is an at most countable index set, with linear part

$$\hat{D}_n(k) = \begin{pmatrix} i\omega_n(k) & 0 \\ 0 & -i\omega_n(k) \end{pmatrix},$$

where the $\omega_n : \mathbb{R} \to \mathbb{R}$ are piece-wise analytic curves, and with the nonlinear operator $\hat{N}_n$ whose k-dependence is also piece-wise analytic. The way to bring the equations into this "natural form" is as follows.

- For the case of the constant coefficient cubic Klein–Gordon equation (1.2), we have the index set $\mathcal{N} = \{1\}$. Thus, the corresponding "system in natural coordinates" (2.5) is two-dimensional, which is consistent with the fact that the dispersion relation has one pair of branches (see left panel of Fig. 2). The transformation to "natural coordinates" is given by $u = [\mathcal{F}\{Q\hat{V}\}]_1$ where $\mathcal{F}$ is the Fourier transform and Q is the 2×2-matrix that diagonalizes the linear part of the equation converted into a first order system.
- For the case of the Maxwell–Lorentz system (2.2)–(2.3), we have the index set $\mathcal{N} = \{1,2\}$ and, so, the corresponding "system in natural coordinates" (2.5) has two two-dimensional subsystems, which is consistent with the fact that the dispersion relation has two pairs of branches (see middle panel of Fig. 2). The transformation to "natural coordinates" is given by $u = [\mathcal{F}\{Q(k)\hat{V}\}]_1$ where $\mathcal{F}$ is the Fourier transform and $Q(k)$ is the 4×4-matrix that diagonalizes the linear part of the equation converted into a first order system.
- For the case of a periodic coefficient nonlinear wave equation we have the index set $\mathcal{N} = \mathbb{N}$ and so the corresponding "system in natural coordinates" (2.5) is a countably infinite collection of two-dimensional subsystems, which corresponds to the fact that the dispersion relation has infinitely many pairs of branches (see right panel of Fig. 2). The transformation to "natural coordinates" is a little more involved consisting of a Bloch transformation followed by an eigenfunction representation and a diagonalizing transformation as we have seen in Section 2.1.

Since the Maxwell–Lorentz system has no quadratic terms in the nonlinearity the proof of an approximation theorem is similar to that of the subsequent approximation Theorem 2. However, before starting the complete program for the Maxwell–Lorentz system one has to get rid of a double zero eigenvalue and a corresponding Jordan block at the Fourier wave number $k = 0$. This Jordan block can be avoided by a suitable choice of variables such that the diagonalization to the "natural coordinates" can be performed, see [6, Section 5].

2.3. Approximation of the system in "natural coordinates"

For our further investigations it will be more convenient to write system (2.1) component-wise as

$$\partial_t \tilde{v}_j(l,t) = i\omega_j(l)\tilde{v}_j(l,t) + \int_{-1/2}^{1/2}\int_{-1/2}^{1/2}\sum_{j_k\in\mathbb{Z}\setminus\{0\}} \tilde{s}^j_{j_1,j_2,j_3}(l, l-l_2-l_3, l_2, l_3) \times \tilde{v}_{j_1}(l-l_2-l_3,t)\tilde{v}_{j_2}(l_2,t)\tilde{v}_{j_3}(l_3,t)\,dl_2\,dl_3, \tag{2.6}$$

for $j\in\mathbb{Z}\setminus\{0\}$ and where $\omega_{-j}(l) = -\omega_j(l)$ and with piece-wise smooth coefficient functions

$$\tilde{s}^j_{j_1,j_2,j_3}(l_1,l_2,l_3,l_4) = -\frac{i}{4\omega_j(l_1)}\big\langle b(\cdot) f_{|j_1|}(l_2,\cdot) f_{|j_2|}(l_3,\cdot) f_{|j_3|}(l_4,\cdot),\, f_{|j|}(l_1,\cdot)\big\rangle. \tag{2.7}$$

As already explained, it is difficult to handle the Bloch transform of the pulse interaction ansatz. It is not even clear at this point what the right modifications to the ansatz (1.4) would be (recall the discussion in Section 1.2). Hence, it is desirable to transform the system back to x-space. Directly transforming back (2.6) is not possible. Hence, before switching back to x-space we will work out an approximation of the operators in (2.6) such that they can be interpreted as the Fourier transform of constant coefficient partial differential operators. Note that, in contrast to the transformations we performed in the last subsection which could be performed regardless of the solution one is interested in, the modifications in this subsection are only justified for the special case of wave packets in NLS form. This is due to the spectral concentration property already illustrated in Section 2.2. Viewed from the perspective of system (2.6), it is clear that an ansatz

$$\tilde{v}_1(l,t) = \tilde{B}\left(\frac{l-l_1}{\varepsilon}, \varepsilon^2 t\right) e^{i(\omega_1(l_1)+\omega_1'(l_1)(l-l_1))t}, \tag{2.8}$$

$\tilde{v}_n = 0$ for $|n|>1$, and $\tilde{v}_{-1}(l) = \overline{\tilde{v}_1(-l)}$ corresponds to a single pulse ansatz for the original system (1.1), i.e.

$$\begin{aligned} u(x,t) &= \frac{1}{\sqrt{2}}\int_{-1/2}^{1/2}\big(\tilde{v}_1(l,t)+\tilde{v}_{-1}(l,t)\big) f_1(l,x) e^{ilx}\,dl \\ &= \frac{1}{\sqrt{2}}\int_{\mathbb{R}}\chi(l)\big(\tilde{v}_1(l,t)+\tilde{v}_{-1}(l,t)\big) f_1(l,x) e^{ilx}\,dl + \mathcal{O}\big(\varepsilon^2\big) \\ &= \varepsilon B\big(\varepsilon\big(x-\omega_1'(l_1)t\big),\varepsilon^2 t\big) f_1(l_1,x) e^{il_1x+i\omega_1(l_1)t} + c.c. + \mathcal{O}\big(\varepsilon^2\big) \end{aligned} \tag{2.9}$$

where $B = \mathcal{F}^{-1}(\chi\tilde{B}(\varepsilon^{-1}(\cdot - l_1)))$ and χ is some cut-off function (see next paragraph). Due to the convolution in system (2.6), an ansatz of the form (2.8) will create terms concentrated at integer multiples of the basic frequency l_1 (in the residual corresponding to (2.6)). Similarly, choosing a two-pulse ansatz will create terms concentrated at linear combinations of the two basic frequencies l_1, l_2.

In summary, the spectral content of pulse solutions is concentrated around the spatial wave number of the carrier wave, such that for the linear dynamics only the form of curves of eigenvalues near the basic wave number l_1 plays a role. Moreover, the convolution in Bloch and Fourier space, respectively, preserves this concentration around integer multiples of the basic spatial wave number l_1.

Applying a cut-off in Bloch space. The first simplifying approximation motivated by the spectral concentration of wave packets is the application of a cut-off function to the kernel of the integrals in (2.6) in order to render the Bloch convolution on the finite interval $[-1/2, 1/2)$ to a Fourier convolution on the real line. To be more precise, the infinite-dimensional system (2.6) in Bloch space is cut off for wave numbers l with $|l| \geqslant 7\max(|l_1|, |l_2|)$ such that the resulting system can be interpreted as infinite-dimensional system in Fourier space. The resulting system has the advantage that its associated inverse Fourier transform representation in x-space is spatially homogeneous with eigenfunctions not depending on the wave number. The cut-off is chosen in such a way that the higher order harmonics are not affected by the cut-off up to the approximation order that we have to consider for our purposes.

Now the starting point of our analysis is system (2.6). For notational simplicity let us assume first that $\max(|l_1|, |l_2|) \leqslant 1/20$. Define the smooth cut-off function χ by $\chi(l) = 1$ for $|l| \leqslant 2/5$, $\chi(l) = 0$ for $|l| \geqslant 9/20$, and $\chi(l) \in [0, 1]$. Then modify (2.6) according to

$$\partial_t \hat{v}_j(l,t) = i\omega_j(l)\hat{v}_j(l,t) + \int_{\mathbb{R}}\int_{\mathbb{R}} \sum_{j_1,j_2,j_3 \in \mathbb{Z}\setminus\{0\}} s^j_{j_1,j_2,j_3}(l, l-l_2-l_3, l_2, l_3) \times \hat{v}_{j_1}(l-l_2-l_3,t)\hat{v}_{j_2}(l_2,t)\hat{v}_{j_3}(l_3,t)\,dl_2\,dl_3, \tag{2.10}$$

with

$$s^j_{j_1,j_2,j_3}(l, l-l_2-l_3, l_2, l_3) = \chi(l)\tilde{s}^j_{j_1,j_2,j_3}(l, l-l_2-l_3, l_2, l_3)\chi(l-l_2-l_3)\chi(l_2)\chi(l_3).$$

System (2.10) is abbreviated subsequently by

$$\partial_t \hat{V}_n(l,t) = \hat{D}_n(l)\hat{V}_n(l,t) + \hat{N}_n\big((\hat{V}_j)_{j\in\mathbb{N}}\big)(l,t), \tag{2.11}$$

where $\hat{D}_n(l) = \tilde{D}_n(l)$ for all $l \in \mathbb{R}$. In general, systems (2.6) and (2.10) display very different dynamics, however, for wave packets in NLS form (2.10) is a good approximation of (2.6) due to the afore mentioned concentration property. The approximation quality can be estimated in terms of the small perturbation parameter $0 < \varepsilon \ll 1$, cf. Lemma 2.6, that measures the concentration.

In more detail, the specific range of l_1, l_2 and the support of χ can be explained as follows. In order to make the residual, i.e. the collection of all terms which remain after inserting the ansatz into the equation, as small as necessary for the description of the pulse interaction, six times the convolution of the basic wave packet has to be taken into account. Hence, only wave numbers smaller than $7\max(|l_1|, |l_2|)$ have to be considered in (2.6). Therefore, what is left out by the cut-off is small. This formal argument is made rigorous with the help of the subsequent Lemma 2.6.

Expansion of multipliers and kernels. We will illustrate the underlying idea by simply considering

$$\partial_t \hat{v}_1(l,t) = i\omega_1(l)\hat{v}_1(l,t) \tag{2.12}$$

with the ansatz

$$\hat{v}_1(l,t) = \tilde{B}(K_1,t), \quad K_1 = \frac{l-k_1}{\varepsilon},$$

with a function $\tilde{B}$ localized in an $\mathcal{O}(\varepsilon)$-neighborhood of (the wave number) $k_1 \in \mathbb{R}$. This transforms (2.12) formally into

$$\partial_t \tilde{B}(K_1,t) = i\left(\sum_{r=0}^{n} \frac{\omega_1^{(r)}(k_1)}{r!}\varepsilon^r K_1^r\right)\tilde{B}(K_1,t) + \mathcal{O}\big(\varepsilon^{n+1}\big)$$

of which inverse Fourier transform (w.r.t. l) yields

$$\varepsilon\partial_t B(X,t)e^{ik_1x} = i\left(\left(\sum_{r=0}^{n} \frac{\omega_1^{(r)}(k_1)}{r!}\varepsilon^{r+1}(-i\partial_X)^r\right)B(X,t)\right)e^{ik_1x} + \mathcal{O}\big(\varepsilon^{n+2}\big), \tag{2.13}$$

with $X := \varepsilon x$ and where $\mathcal{F}^{-1}(\tilde{B}(\cdot,t)) = B$.

Remark 2.4. Motivated by the multiple scales arising in Eq. (2.13) we could set now

$$B = \mathcal{B}\big(X,t,\varepsilon t,\varepsilon^2 t,\ldots\big)$$

which gives directly the wave packet ansatz

$$\mathcal{B}\big(X,t,\varepsilon t,\varepsilon^2 t,\ldots\big) = A\big(X+\varepsilon\omega_1'(k_1)t,\varepsilon^2 t\big)e^{i\omega_1(k_1)t} + \mathcal{O}\big(\varepsilon^3 t\big).$$

The above idea can obviously be carried over to multiple wave packets and to the nonlinearity by using the fact that the convolution spreads the localization around integer multiples of the initially chosen (wave) number k_1 (or resp. around linear integer combinations of k_1 and k_2 in case of a two-pulse ansatz). Hence, one only has to keep track of the different terms created by the convolution and expand the kernel functions $s^j_{j_1j_2j_3}$ in (2.10) correspondingly. We skip the details at this point and simply display the terms in the next section where we compute solvability conditions for a two-pulse ansatz.

Estimates for the approximation through cut-off and expansion. As already alluded to, in order to stay as close as possible to the procedure from [5], we want to employ an extended two-pulse ansatz to an equation in x-space. Hence, after performing the above approximations to each term in (2.6) accordingly we perform inverse Fourier transform. The approximations of the terms in (2.6) can be justified with the subsequent fundamental lemma.

Definition 2.5. We define the space

$$L^2(s) = \big\{u:\mathbb{R}\to\mathbb{C}\;\big|\;\|\hat{u}\|_{L^2(s)} < \infty\big\}$$

equipped with the norm

$$\|\hat{u}\|_{L^2(s)} = \left(\int \big|\hat{u}(k)\big|^2\big(1+k^2\big)^s\,dk\right)^{1/2}.$$

Lemma 2.6. *Fix $s \leqslant s_B - (n+1)$. Then let $\mu:\mathbb{R}\to\mathbb{R}$ be an $(n+1)$-times differentiable function satisfying $|\mu(k)| \leqslant C|k-k_1|^{n+1}$ and let $v(x) = \varepsilon B(\varepsilon x)e^{ik_1x}$, with $B \in H^{s_B}$. Then there exists a C such that for all $0 \leqslant \varepsilon \leqslant 1$ we have*

$$\left\|\big(\mathcal{F}^{-1}\mu\mathcal{F}\big)v - \left(\mathcal{F}^{-1}\left(\sum_{m=0}^{n}\frac{\mu^{(m)}(k_1)}{m!}(\cdot-k_1)^m\right)\mathcal{F}\right)v\right\|_{H^s} \leqslant C\varepsilon^{n+1+1/2}\|B\|_{H^{s_B}}. \tag{2.14}$$

Proof. The proof merely uses Taylor expansion and that $\mathcal{F}$ is an isometry between H^s and $L^2(s)$. In detail,

$$\begin{aligned}
&\left\| \left(\mathcal{F}^{-1}\mu\mathcal{F}\right)v - \left(\mathcal{F}^{-1}\left(\sum_{m=0}^{n}\frac{\mu^{(m)}(k_1)}{m!}(\cdot-k_1)^m\right)\mathcal{F}\right)v \right\|_{H^s} \\
&\quad\leqslant C\varepsilon^{1/2+(n+1)} \sup_{K_1\in\mathbb{R}}\left|\frac{(1+K_1^2)^{s/2}K_1^{n+1}}{(1+K_1^2)^{s_B/2}}\right| \|\hat{B}\|_{L^2(s_B)} \\
&\quad\leqslant C\varepsilon^{1/2+(n+1)}\|B\|_{H^{s_B}}
\end{aligned}$$

if $s < s_B - (n+1)$. □

Hence, the error which is made by replacing the linear pseudo differential operator $\mathcal{F}^{-1}\hat{D}_1\mathcal{F}$ by a suitable linear partial differential operator $\underline{D}_{k_1}$ is small if both are applied to a wave packet $\mathcal{F}^{-1}\hat{V}^{an}$ in NLS scaling with basic wave number k_1. Another corollary is the smallness of the difference between the original system and the cut-off system.

Corollary 2.7. *Fix $s \leqslant s_B - (n+1)$ and $|k_1| < b$. Then let $v(x) = \varepsilon B(\varepsilon x)e^{ik_1x}$, with $B \in H^{s_B}$. Then there exists a C such that for all $0 \leqslant \varepsilon \leqslant 1$ we have*

$$\left\| \left(\mathcal{F}^{-1}\chi_{[-b,b]}\mathcal{F}\right)v - \left(\mathcal{F}^{-1}\left(\sum_{m=0}^{n}\frac{\mu^{(m)}(k_1)}{m!}(\cdot-k_1)^m\right)\mathcal{F}\right)v \right\|_{H^s} \leqslant C\varepsilon^{n+1+1/2}\|B\|_{H^{s_B}}. \quad (2.15)$$

Proof. Use Lemma 2.6 with $\mu = \chi_{[-b,b]}$. We have $|\mu(k)| \leqslant C|k-k_1|^{n+1}$ for all $n \in \mathbb{N}$ which implies the assertion. □

Remark 2.8. The statements of Lemma 2.6 and Corollary 2.7 hold for multilinear mappings, too.

3. Derivation of the extended modulation system

The procedure for the derivation of the extended modulation system is now very similar to the one made in [5, Section 3]. There are a number of minor changes which however makes it necessary to redo these calculations.

For the derivation of the extended modulation system we transfer (2.10) into x-space and make the extended two-pulse ansatz V^{an} defined by

$$v_1^{an}(x,t) = \sum_{j=1,2}\sum_{r=1}^{3}\varepsilon^r A_j^{(r)}(X_j,T)e^{iY_j} + M_{mixed,1}, \quad (3.1)$$

$$v_1^{an}(x,t) = \overline{v_{-1}^{an}(x,t)}, \quad (3.2)$$

$$v_j^{an}(x,t) = M_{mixed,j} \quad \text{for } j \in \mathbb{Z}\setminus\{-1,1\}, \quad (3.3)$$

$$X_j = X + \varepsilon\omega_1'(k_j)t + \varepsilon^2\Psi_j^{(1)}\left(X+\varepsilon\omega_1'(k_{3-j})t, T\right), \quad (3.4)$$

$$Y_j = k_jx + \omega_1(k_j)t + \sum_{l=1,2}\varepsilon^l\Omega_j^{(l)}\left(X+\varepsilon\omega_1'(k_{3-j})t, T\right), \quad (3.5)$$

$$X = \varepsilon x, \qquad T = \varepsilon^2 t. \quad (3.6)$$

The mixed and higher order harmonic terms $M_{mixed,j}$ satisfy

$$M_{mixed,j} = \mathcal{O}(\varepsilon^3).$$

Remark 3.1. There are essentially two possibilities. The two wave packets belong either to the same curve of eigenvalues, say ω_n, or to different curves of eigenvalues, say ω_{n_1} and ω_{n_2}. In the first case the consideration of a two-dimensional subsystem is sufficient, in the second case the consideration of a two- (if the two curves are given by ω_n and ω_{-n}) or a four-dimensional subsystem is sufficient. Both cases can be treated by the same procedure. Therefore, we restrict ourselves for notational simplicity and clarity to the first case, namely that the basic spatial wave numbers k_1 and k_2 belong to the same curves of eigenvalues, say ω_1. In the second case we could restrict to the curves of eigenvalues ω_1 and ω_2 or to ω_1 and ω_{-1}. The general situation can always be brought to one of these cases by renumbering the variables.

As already explained in the last section, we would like to stay as close as possible to the pulse interaction analysis from [5] and, hence, do the derivation of modulation equations in x-space. To that end, we will work with an approximation of (2.6) obtained by the procedure illustrated in the previous section adapted to the two-pulse case. We denote the n-th component of the residual (corresponding to the inverse Fourier transformed approximation of (2.6)) by

$$\mathrm{Res}_n(V)(x,t) = -\partial_t V_n(x,t) + D_n(\partial_x)V_n(x,t) + N_n\big((V_j)_{j\in\mathbb{N}}\big)(x,t), \tag{3.7}$$

w.r.t. $\varepsilon^l e^{i(mY_1+nY_2)}$, i.e.,

$$\widehat{\mathrm{Res}}_j(V) = \sum_{j_1,j_2,j_3} \varepsilon^{j_1}\,\mathrm{Res}^j_{j_1,j_2,j_3}(V)e^{i(j_2Y_1+j_3Y_2)} \tag{3.8}$$

which is an implicit definition for the terms $\mathrm{Res}^j_{j_1,j_2,j_3}(V)$. Plugging in the ansatz V^{an} leads to a number of conditions in order to make the residual as small as possible, in particular to nonlinear Schrödinger equations for the $A^{(1)}_j$.

Inserting this ansatz into the nonlinear terms of (3.7) gives terms which are in Fourier space of the form

$$\varepsilon^\gamma \int_{\mathbb{R}}\int_{\mathbb{R}} s(k,k-l-m,l,m)\hat{B}_1\left(\frac{k-l-m-k_1}{\varepsilon}\right)\hat{B}_2\left(\frac{m-k_2}{\varepsilon}\right)\hat{B}_3\left(\frac{l-k_3}{\varepsilon}\right)dm\,dl \tag{3.9}$$

with a piece-wise smooth kernel $s = s^j_{j_1,j_2,j_3}$ where $j_1, j_2, j_3 \in \{-1,1\}$, $\gamma \in \mathbb{N}$, and $\hat{B}_\nu \in \{A^{(1)}_1,\dots,\overline{A^{(3)}_2}\}$ for $\nu \in \{1,2,3\}$. As already explained in x-space this expression corresponds to a multilinear pseudo differential operator. Similar to the procedure in Section 2 we would like to replace this operator by a multilinear partial differential operator.

Similar as above the concentration of the ansatz at various wave numbers causes an expansion of the kernel function s at these wave numbers. Therefore, (3.9) has an expansion

$$\varepsilon^{\gamma+2}s(k_1+k_2+k_3,k_1,k_2,k_3)\left(\int_{\mathbb{R}}\int_{\mathbb{R}}\hat{B}_1(K_1-\underline{M}-\underline{L})\hat{B}_2(\underline{M})\hat{B}_3(\underline{L})\,d\underline{M}\,d\underline{L}\right)+\mathcal{O}(\varepsilon^{\gamma+3}).$$

For our purposes we have to consider terms up to order $\mathcal{O}(\varepsilon^4)$ in Fourier space, i.e., we have to consider for instance the expansion $I_1+I_2+I_3$ for those terms where $\hat{B}_\nu \in \{A^{(1)}_1,\dots,\overline{A^{(1)}_2}\}$ for $\nu \in \{1,2,3\}$, where

$$I_1 = \varepsilon^2 s(k_1+k_2+k_3,k_1,k_2,k_3)\Big(\int_{\mathbb{R}}\int_{\mathbb{R}} \hat{B}_1(K_1-\underline{M}-\underline{L})\hat{B}_2(\underline{M})\hat{B}_3(\underline{L})\,d\underline{M}\,d\underline{L}\Big),$$

$$\begin{aligned} I_2 = {} & \varepsilon^3 \partial_1 s(k_1+k_2+k_3,k_1,k_2,k_3)\Big(\int_{\mathbb{R}}\int_{\mathbb{R}} K_1\hat{B}_1(K_1-\underline{M}-\underline{L})\hat{B}_2(\underline{M})\hat{B}_3(\underline{L})\,d\underline{M}\,d\underline{L}\Big) \\ & + \varepsilon^3 \partial_2 s(k_1+k_2+k_3,k_1,k_2,k_3) \\ & \times\Big(\int_{\mathbb{R}}\int_{\mathbb{R}} (K_1-\underline{M}-\underline{L})\hat{B}_1(K_1-\underline{M}-\underline{L})\hat{B}_2(\underline{M})\hat{B}_3(\underline{L})\,d\underline{M}\,d\underline{L}\Big) \\ & + \varepsilon^3 \partial_3 s(k_1+k_2+k_3,k_1,k_2,k_3)\Big(\int_{\mathbb{R}}\int_{\mathbb{R}} \hat{B}_1(K_1-\underline{M}-\underline{L})\underline{M}\hat{B}_2(\underline{M})\hat{B}_3(\underline{L})\,d\underline{M}\,d\underline{L}\Big) \\ & + \varepsilon^3 \partial_4 s(k_1+k_2+k_3,k_1,k_2,k_3)\Big(\int_{\mathbb{R}}\int_{\mathbb{R}} \hat{B}_1(K_1-\underline{M}-\underline{L})\hat{B}_2(\underline{M})\underline{L}\hat{B}_3(\underline{L})\,d\underline{M}\,d\underline{L}\Big), \end{aligned}$$

$$\begin{aligned} I_3 = {} & \varepsilon^4 \partial_1^2 s(k_1+k_2+k_3,k_1,k_2,k_3)\Big(\int_{\mathbb{R}}\int_{\mathbb{R}} K_1^2\hat{B}_1(K_1-\underline{M}-\underline{L})\hat{B}_2(\underline{M})\hat{B}_3(\underline{L})\,d\underline{M}\,d\underline{L}\Big) \\ & \vdots \\ & + \varepsilon^4 \partial_4^2 s(k_1+k_2+k_3,k_1,k_2,k_3)\Big(\int_{\mathbb{R}}\int_{\mathbb{R}} \hat{B}_1(K_1-\underline{M}-\underline{L})\hat{B}_2(\underline{M})\underline{L}^2\hat{B}_3(\underline{L})\,d\underline{M}\,d\underline{L}\Big). \end{aligned}$$

Additionally, we have to consider the expansion $\varepsilon(I_1+I_2)$ for those terms where for two indices $\nu\in\{1,2,3\}$ we have $\hat{B}_\nu\in\{A_1^{(1)},\dots,\overline{A_2^{(1)}}\}$ and where for one index $\nu\in\{1,2,3\}$ we have $\hat{B}_\nu\in\{A_1^{(2)},\dots,\overline{A_2^{(2)}}\}$. Finally, we have to consider the expansion $\varepsilon^2 I_1$ for those terms where for two indices $\nu\in\{1,2,3\}$ we have $\hat{B}_\nu\in\{A_1^{(1)},\dots,\overline{A_2^{(1)}}\}$ and where for one index $\nu\in\{1,2,3\}$ we have $\hat{B}_\nu\in\{A_1^{(3)},\dots,\overline{A_2^{(3)}}\}$ and for those terms where for two indices $\nu\in\{1,2,3\}$ we have $\hat{B}_\nu\in\{A_1^{(2)},\dots,\overline{A_2^{(2)}}\}$ and where for one index $\nu\in\{1,2,3\}$ we have $\hat{B}_\nu\in\{A_1^{(1)},\dots,\overline{A_2^{(1)}}\}$.

The mixed terms: By the nonlinearity various mixed terms, as for instance

$$\begin{aligned} & \varepsilon^{r_1+r_2+r_3} s^j_{j_1,j_2,j_3}(k_{j_1}+k_{j_2}+k_{j_3},k_{j_1},k_{j_2},k_{j_3}) \\ & \quad\times A_{j_1}^{(r_1)}(X_{j_1},T)A_{j_2}^{(r_2)}(X_{j_2},T)A_{j_3}^{(r_3)}(X_{j_3},T)e^{i(Y_{j_1}+Y_{j_2}+Y_{j_3})} \end{aligned}$$

or for instance

$$\begin{aligned} & \varepsilon^{r_1+r_2+r_3+1} \partial_2 s^j_{j_1,j_2,j_3}(k_{j_1}+k_{j_2}+k_{j_3},k_{j_1},k_{j_2},k_{j_3}) \\ & \quad\times A_{j_1}^{(r_1)}(X_{j_1},T)\partial_1 A_{j_2}^{(r_2)}(X_{j_2},T)A_{j_3}^{(r_3)}(X_{j_3},T)e^{i(Y_{j_1}+Y_{j_2}+Y_{j_3})} \end{aligned}$$

in the j-th equation are created. All terms which are of order $\mathcal{O}(\varepsilon^3)$, $\mathcal{O}(\varepsilon^4)$, or $\mathcal{O}(\varepsilon^5)$ have to be eliminated. This will be done by the terms contained in $M_{mixed,j}$ present in the above ansatz V^{an}. In order to eliminate the first of the above terms we add

$$\varepsilon^{r_1+r_2+r_3} M^{r_1,r_2,r_3}_{j_1,j_2,j_3}(X,\varepsilon t) e^{i(Y_{j_1}+Y_{j_2}+Y_{j_3})}$$

to $M_{mixed,j}$ with

$$\begin{aligned} M^{r_1,r_2,r_3}_{j_1,j_2,j_3}(X,\varepsilon t) = {} & \big(\big(\omega_1(k_{j_1})+\omega_1(k_{j_2})+\omega_1(k_{j_3})\big)^2 - \omega_j(k_{j_1}+k_{j_2}+k_{j_3})^2\big)^{-1} \\ & \times s^j_{j_1,j_2,j_3}(k_{j_1}+k_{j_2}+k_{j_3}, k_{j_1}, k_{j_2}, k_{j_3}) \\ & \times A^{(r_1)}_{j_1}(X_{j_1},T) A^{(r_2)}_{j_2}(X_{j_2},T) A^{(r_3)}_{j_3}(X_{j_3},T). \end{aligned}$$

Similarly all other mixed terms can be eliminated such that after a correct choice we have

$$\mathrm{Res}^j_{j_1,j_2,j_3} = 0$$

for $j_1 \leqslant 5$ if $|j| \geqslant 2$ or if

$$|j| = 1, \quad (j_2,j_3) \notin \big\{(1,0),(-1,0),(0,-1),(0,1)\big\}.$$

Thus, we can concentrate on the remaining terms of the residual.

The main terms: Before we start we remark that due to

$$\begin{aligned} & \partial_X\big(A^{(r)}_j\big(X_j+\varepsilon^2\Psi^{(1)}_j\big(X+\varepsilon\omega_1'(k_{3-j})t,T\big),T\big) e^{ik_jx+i\omega_1(k_j)t+i\sum_{l=1,2}\varepsilon^l\Omega^{(l)}_j(X+\varepsilon\omega_1'(k_{3-j})t,T)}\big) \\ & = \partial_X\big(A^{(r)}_j(X_j,T)e^{ik_jx+i\omega_1(k_j)t}\big) + \mathcal{O}\big(\varepsilon^2\big), \end{aligned}$$

there will be no terms which are nonlinear w.r.t. $\Omega^{(l)}_j$ or $\Psi^{(1)}_j$ in the relevant residual terms.

Remark 1. We remark that the kernel function s^1_{111} evaluated at $(k,k,l,-l)$ is purely imaginary (which can be read off directly from its definition (2.7)). As a consequence the sums $\partial_1 s^1_{111}(k,k,l,-l) + \partial_2 s^1_{111}(k,k,l,-l)$ and $\partial_3 s^1_{111}(k,k,l,-l) + \partial_4 s^1_{111}(k,k,l,-l)$ are also purely imaginary.

Using the notation $\mathrm{Res}^1_{l,m,n}$ from (3.8) for the coefficients of $\varepsilon^l e^{i(mY_1+nY_2)}$ we find the subsequent hierarchy of equations.

- At εe^{iY_1} we find $\mathrm{Res}^1_{1,1,0} = 0$ due to the direct inclusion of $\omega_1(k_1)$ in the ansatz.
- At $\varepsilon^2 e^{iY_1}$ we find $\mathrm{Res}^1_{2,1,0} = 0$ due to the direct inclusion of $\omega_1'(k_1)$ in the ansatz.
- At $\varepsilon^3 e^{iY_1}$ we find

$$\mathrm{Res}^1_{3,1,0} = t_{31} + \tilde{t}_{32} A^{(1)}_1$$

with

$$\begin{aligned} t_{31} = {} & -\partial_2 A^{(1)}_1(\underline{X}_1,T) - i\big(\omega_1''(k_1)/2\big)\partial_1^2 A^{(1)}_1(\underline{X}_1,T) \\ & + 3s^1_{111}(k_1,k_1,k_1,-k_1)\big|A^{(1)}_1(\underline{X}_1,T)\big|^2 A^{(1)}_1(\underline{X}_1,T), \\ \tilde{t}_{32} = {} & i\big(\omega_1'(k_1)-\omega_1'(k_2)\big)\partial_1\Omega^{(1)}_1(\underline{X}_2,T) \\ & + 6s^1_{111}(k_1,k_1,k_2,-k_2)\big|A^{(1)}_2(\underline{X}_2,T)\big|^2. \end{aligned}$$

Then the condition $t_{31}=0$ yields the NLS equation

$$\partial_2 A_1^{(1)}(\underline{X}_1,T)=-i\big(\omega_1''(k_1)/2\big)\partial_1^2 A_1^{(1)}(\underline{X}_1,T)+3s_{111}^1(k_1,k_1,k_1,-k_1)\big|A_1^{(1)}(\underline{X}_1,T)\big|^2A_1^{(1)}(\underline{X}_1,T),\tag{3.10}$$

where $s_{111}^1(k_1,k_1,k_1,-k_1)\in i\mathbb{R}$ and the condition $\tilde{t}_{32}=0$ yields the phase shift formula

$$\Omega_1^{(1)}(\underline{X}_2,T)=\frac{6s_{111}^1(k_1,k_1,k_2,-k_2)}{i(\omega_1'(k_2)-\omega_1'(k_1))}\int^{X_2}\big|A_2^{(1)}(\zeta,T)\big|^2\,\mathrm{d}\zeta,\tag{3.11}$$

so $\Omega_1^{(1)}$ is a real quantity since $s_{111}^1(k_1,k_1,k_2,-k_2)\in i\mathbb{R}$ and therefore a pure phase correction.

- At $\varepsilon^4e^{iY_1}$ we find

$$\begin{aligned}\mathrm{Res}_{4,1,0}^1=&-\partial_2A_1^{(2)}(\underline{X}_1,T)-i\big(\omega_1''(k_1)/2\big)\partial_1^2A_1^{(2)}(\underline{X}_1,T)\\&+t_{41}+\tilde{t}_{42}A_1^{(1)}(\underline{X}_1,T)+\tilde{t}_{43}\partial_1A_1^{(1)}(\underline{X}_1,T)\end{aligned}$$

where

$$\begin{aligned}
t_{41a}&=s_{111}^1(k_1,k_1,k_1,-k_1)\big(6A_1^{(2)}(\underline{X}_1,T)\overline{A_1^{(1)}}(\underline{X}_1,T)+3\overline{A_1^{(2)}}(\underline{X}_1,T)A_1^{(1)}(\underline{X}_1,T)\big)A_1^{(1)}(\underline{X}_1,T),\\
t_{41b}&=-3i\,\partial_1s_{111}^1(k_1,k_1,k_1,-k_1)\,\partial_1\big(A_1^{(1)}(\underline{X}_1,T)\overline{A_1^{(1)}}(\underline{X}_1,T)A_1^{(1)}(\underline{X}_1,T)\big),\\
t_{41c}&=-6i\,\partial_2s_{111}^1(k_1,k_1,k_1,-k_1)\big(\partial_1A_1^{(1)}\big)(\underline{X}_1,T)\overline{A_1^{(1)}}(\underline{X}_1,T)A_1^{(1)}(\underline{X}_1,T),\\
t_{41d}&=-6i\,\partial_3s_{111}^1(k_1,k_1,k_1,-k_1)A_1^{(1)}(\underline{X}_1,T)\overline{A_1^{(1)}}(\underline{X}_1,T)\big(\partial_1A_1^{(1)}\big)(\underline{X}_1,T),\\
t_{41e}&=-6i\,\partial_4s_{111}^1(k_1,k_1,k_1,-k_1)A_1^{(1)}(\underline{X}_1,T)\big(\partial_1\overline{A_1^{(1)}}\big)(\underline{X}_1,T)A_1^{(1)}(\underline{X}_1,T),\\
t_{41f}&=i\big(\omega_1'''(k_1)/6\big)\partial_1^3A_1^{(1)}(\underline{X}_1,T),\\
\tilde{t}_{42a}&=6s_{111}^1(k_1,k_1,k_2,-k_2)\big(A_2^{(2)}(\underline{X}_2,T)\overline{A_2^{(1)}}(\underline{X}_2,T)+\overline{A_2^{(2)}}(\underline{X}_2,T)A_2^{(1)}(\underline{X}_2,T)\big),\\
\tilde{t}_{42b}&=-3i\,\partial_1s_{111}^1(k_1,k_1,k_2,-k_2)\,\partial_1\big(A_2^{(1)}(\underline{X}_2,T)\overline{A_2^{(1)}}(\underline{X}_2,T)\big),\\
\tilde{t}_{43a}&=-6i\,\partial_1s_{111}^1(k_1,k_1,k_2,-k_2)\big(A_2^{(1)}(\underline{X}_2,T)\overline{A_2^{(1)}}(\underline{X}_2,T)\big),\\
\tilde{t}_{43b}&=-6i\,\partial_2s_{111}^1(k_1,k_1,k_2,-k_2)\big(A_2^{(1)}(\underline{X}_2,T)\overline{A_2^{(1)}}(\underline{X}_2,T)\big),\\
\tilde{t}_{42c}&=-6i\,\partial_3s_{111}^1(k_1,k_1,k_2,-k_2)\big(\partial_1A_2^{(1)}\big)(\underline{X}_2,T)\overline{A_2^{(1)}}(\underline{X}_2,T),\\
\tilde{t}_{42d}&=-6i\,\partial_4s_{111}^1(k_1,k_1,k_2,-k_2)A_2^{(1)}(\underline{X}_2,T)\big(\partial_1\overline{A_2^{(1)}}\big)(\underline{X}_2,T),\\
\tilde{t}_{42e}&=i\partial_2\Omega_1^{(1)}(\underline{X}_2,T),\\
\tilde{t}_{43c}&=\big(\big(\omega_1''(k_1)/2\big)\partial_1\Omega_1^{(1)}(\underline{X}_2,T)+\big(\omega_1'(k_1)-\omega_1'(k_2)\big)\partial_1\Psi_1^{(1)}(\underline{X}_2,T)\big),\\
\tilde{t}_{42f}&=\big(\big(\omega_1''(k_1)/2\big)\partial_1^2\Omega_1^{(1)}(\underline{X}_2,T)+i\big(\omega_1'(k_1)-\omega_1'(k_2)\big)\partial_1\Omega_1^{(2)}(\underline{X}_2,T)\big).
\end{aligned}$$

We can now choose $A_1^{(2)}$ to solve the linear inhomogeneous evolution equation

$$\partial_2 A_1^{(2)}(\underline{X}_1, T) = i\big(\omega_1''(k_1)/2\big)\partial_1^2 A_1^{(2)}(\underline{X}_1, T) + t_{41}. \tag{3.12}$$

Here, no coupling with terms involving $A_2^{(r)}$-variables occurs such that $A_1^{(2)}$ describes internal dynamics of a single pulse.
The remaining terms $\tilde{t}_{42}A_1^{(1)} + \tilde{t}_{43}\partial_1 A_1^{(1)}$ will be used to derive solvability conditions for $\Psi_1^{(1)}$, $\Omega_1^{(2)}$:

First we observe that the only term involving $\Psi_1^{(1)}$ is proportional to $\partial_1 A_1^{(1)}$. Therefore, we collected all terms that are also proportional to $\partial_1 A_1^{(1)}$ and use them to derive an equation for $\Psi_1^{(1)}$, i.e., we set

$$\begin{aligned}\tilde{t}_{43} = {} & \big(\big(\omega_1''(k_1)/2\big)\partial_1\Omega_1^{(1)}(\underline{X}_2, T) + \big(\omega_1'(k_1) - \omega_1'(k_2)\big)\partial_1\Psi_1^{(1)}(\underline{X}_2, T)\big)\\ & - 6i\,\partial_1 s_{111}^1(k_1, k_1, k_2, -k_2)\big(A_2^{(1)}(\underline{X}_2, T)\overline{A_2^{(1)}}(\underline{X}_2, T)\big)\\ & - 6i\,\partial_2 s_{111}^1(k_1, k_1, k_2, -k_2)\big(A_2^{(1)}(\underline{X}_2, T)\overline{A_2^{(1)}}(\underline{X}_2, T)\big) = 0\end{aligned}$$

which, together with (3.11), gives the envelope shift formula

$$\Psi_1^{(1)}(\underline{X}_2, T) = C(k_1, k_2)\int^{\underline{X}_2} \big|A_2^{(1)}(\zeta, T)\big|^2\,\mathrm{d}\zeta, \tag{3.13}$$

where we refrained from explicitly displaying the prefactor $C(k_1, k_2)$ for better readability and to emphasize the structure. Note that $C(k_1, k_2) \in \mathbb{R}$ again due to Remark 1.
Now all the remaining terms, i.e. $\tilde{t}_{42}A_1^{(1)}$, are proportional to $A_1^{(1)}$ and are perfectly suited to derive an equation for $\Omega_1^{(2)}$ by setting

$$\begin{aligned}\tilde{t}_{42} = {} & 6s_{111}^1(k_1, k_1, k_2, -k_2)\big(A_2^{(2)}(\underline{X}_2, T)\overline{A_2^{(1)}}(\underline{X}_2, T) + \overline{A_2^{(2)}}(\underline{X}_2, T)A_2^{(1)}(\underline{X}_2, T)\big)\\ & - 3i\,\partial_1 s_{111}^1(k_1, k_1, k_2, -k_2)\,\partial_1\big(A_2^{(1)}(\underline{X}_2, T)\overline{A_2^{(1)}}(\underline{X}_2, T)\big)\\ & - 6i\,\partial_3 s_{111}^1(k_1, k_1, k_2, -k_2)\big(\partial_1 A_2^{(1)}\big)(\underline{X}_2, T)\overline{A_2^{(1)}}(\underline{X}_2, T)\\ & - 6i\,\partial_4 s_{111}^1(k_1, k_1, k_2, -k_2)A_2^{(1)}(\underline{X}_2, T)\big(\partial_1\overline{A_2^{(1)}}\big)(\underline{X}_2, T)\\ & + i\partial_2\Omega_1^{(1)}(\underline{X}_2, T)\\ & + \big(\big(\omega_1''(k_1)/2\big)\partial_1^2\Omega_1^{(1)}(\underline{X}_2, T) + i\big(\omega_1'(k_1) - \omega_1'(k_2)\big)\partial_1\Omega_1^{(2)}(\underline{X}_2, T)\big) = 0.\end{aligned} \tag{3.14}$$

The quantity $\Omega_1^{(2)}$ can now be interpreted as follows. Its real part is a second order correction to the phase shift, whereas its imaginary part gives a correction of the amplitude. We refrain from explicitly displaying the rather lengthy expression for the real part of $\Omega_1^{(2)}$ and only note that it is pure integration of spatially localized terms. The imaginary part of $\Omega_1^{(2)}$ satisfies

$$
\begin{aligned}
&-3\,\mathrm{Im}\big(i\,\partial_1 s^1_{111}(k_1,k_1,k_2,-k_2)\,\big(A_2^{(1)}(\underline{X}_2,T)\overline{A_2^{(1)}}(\underline{X}_2,T)\big)\big)\\
&\quad-3i\,\big(\partial_3 s^1_{111}(k_1,k_1,k_2,-k_2)+\partial_4 s^1_{111}(k_1,k_1,k_2,-k_2)\big)\big(A_2^{(1)}\big)(\underline{X}_2,T)\overline{A_2^{(1)}}(\underline{X}_2,T)\\
&\quad+\big(\big(\omega_1''(k_1)/2\big)\partial_1\Omega_1^{(1)}(\underline{X}_2,T)+\big(\omega_1'(k_1)-\omega_1'(k_2)\big)\,\mathrm{Im}\,\Omega_1^{(2)}(\underline{X}_2,T)\big)=0
\end{aligned}
$$

where we used again Remark 1. Note that $\tilde{t}_{42c}$ and $\tilde{t}_{42d}$ contribute to the real part of $\Omega_1^{(2)}$, too. Hence $\mathrm{Im}\,\Omega_1^{(2)}$ is spatially localized and so the induced correction is small w.r.t. ε except during the collision of spatially localized wave packets. Since $\mathrm{Im}\,\Omega_1^{(2)}$, $\mathrm{Im}\,\Omega_2^{(2)}$ are supposed to describe interaction dynamics we may assume that $\Omega_1^{(2)}=\Omega_2^{(2)}=0$ initially. Moreover, due to the fact that $\mathrm{Im}\,\Omega_1^{(2)}$ and $\mathrm{Im}\,\Omega_2^{(2)}$ turn out to be spatially localized in the region of interaction, $\mathrm{Im}\,\Omega_1^{(2)}$ and $\mathrm{Im}\,\Omega_2^{(2)}$ play no role for the envelope shift, cf. Section 4.

- For the purpose we have in mind at $\varepsilon^5 e^{iY_1}$ we do not have to sort terms as carefully as before. We find

$$
\mathrm{Res}^1_{5,1,0}=\partial_2 A_1^{(3)}(\underline{X}_1,T)+i\big(\omega_1''(k_1)\big)\partial_1^2 A_1^{(3)}(\underline{X}_1,T)+t_{51}+t_{52}
$$

where t_{51} contains all terms which solely depend on $\underline{X}_1$ and where t_{52} contains all interaction terms, i.e. all terms which are products of functions of which at least one depends on $\underline{X}_1$ and at least one depends on $\underline{X}_2$. For the same arguments as before there are no terms present which solely depend on $\underline{X}_2$. We choose $A_1^{(3)}$ to satisfy

$$
\partial_2 A_1^{(3)}(\underline{X}_1,T)+i\big(\omega_1''(k_1)\big)\partial_1^2 A_1^{(3)}(\underline{X}_1,T)+t_{51}=0 \tag{3.15}
$$

where t_{51} is at most linear in $A_1^{(3)}$ and $\overline{A_1^{(3)}}$. Moreover, t_{51} only depends on $A_1^{(1)}$, $A_1^{(2)}$, $\Omega_1^{(1)}$, $\Psi_1^{(1)}$, $\Omega_1^{(2)}$ and the lowest order mixed terms which can be expressed again in terms of $A_1^{(1)}$ and $A_1^{(2)}$. Hence, $A_1^{(3)}$ describes pure internal dynamics.
We will demonstrate in the next section an estimate on the interaction terms t_{52} allowing to shift them effectively to the next order in the residual.

Finally we choose $A_2^{(1)}$, $A_2^{(2)}$, $A_2^{(3)}$, $\Omega_2^{(1)}$, $\Omega_2^{(2)}$, and $\Psi_2^{(1)}$ to satisfy the counterparts to (3.10)–(3.15).

4. Estimates for the residual in "natural coordinates"

The goal of this section is the proof of

Theorem 4.1. *Let $s\geqslant 2$, $m\geqslant 2$, $s_A\geqslant s+10$, $l_1\neq l_2$, $l_1,l_2>0$, and let $A_1^{(1)}|_{T=0}, A_2^{(1)}|_{T=0}\in H^{s_A}(m)\cap H^{s_A+m}(0)$. Then for all $T_0>0$ there exist $\varepsilon_0>0$, $C>0$ such that for all $\varepsilon\in(0,\varepsilon_0)$ we have*

$$
\sup_{t\in[0,T_0/\varepsilon^2]}\big\|\mathrm{Res}(V_{an})\big\|_{H^s(\mathbb{R},\ell^2(s))}\leqslant C\varepsilon^{11/2},
$$

with V_{an} given by (3.1)–(3.6), where $\mathrm{Res}(V)=(\mathrm{Res}_n(V))_{n\in\mathbb{Z}\setminus\{0\}}$ and

$$
\mathrm{Res}_n(V)(x,t)=-\partial_t V_n(x,t)+D_n(\partial_x)V_n(x,t)+N_n\big((V_j)_{j\in\mathbb{N}}\big)(x,t).
$$

The difference between the exponents of the formal error $\mathcal{O}(\varepsilon^6)$ and $\mathcal{O}(\varepsilon^{11/2})$ in the lemma follows from the scaling properties of the L^2-norm. The weighted spaces $H^s(m)$ are used to describe

analytically the condition that the wave packets are spatially localized. This is needed to estimate the interaction terms like for instance t_{52}. The loss of regularity from s_A to s is explained below.

In order to prove Theorem 4.1, we prove in Lemma 4.4 the $\mathcal{O}(1)$-boundedness of $\Omega_1^{(1)}$, $\Omega_2^{(1)}$, $\operatorname{Re}\Omega_1^{(2)}$ and $\operatorname{Re}\Omega_2^{(2)}$ in L^∞ and that $\operatorname{Im}\Omega_1^{(2)}$ and $\operatorname{Im}\Omega_2^{(2)}$ are $\mathcal{O}(1)$-bounded in $H^s(m)$. Thus, for instance, $|\operatorname{Im}\Omega_1^{(2)}(X_2,T)| \leqslant C/(1+\varepsilon|x-c_2t|)^m$ due to Sobolev's embedding theorem for $s > 1/2$. For the same reason we have $|A_1^{(j)}(X_1,T)| \leqslant C/(1+\varepsilon|x-c_1t|)^m$ and so, for large t, i.e. for $t > 1/\varepsilon$,

$$\big|A_1^{(j)}(X_1,T)\big|\operatorname{Im}\Omega_1^{(2)}(X_2,T) = \mathcal{O}\big((\varepsilon t)^{-m}\big).$$

In order to have for instance $|A_1^{(j)}(X_1,T)|\operatorname{Im}\Omega_1^{(2)}(X_2,T) = o(\varepsilon)$, except during interaction we require $(\varepsilon t)^{-m} = \mathcal{O}(\varepsilon^{1+\delta m})$ with $\delta > 0$ arbitrary small but fixed. This yields $t \sim \varepsilon^{-(1+1/m+\delta)} \ll \varepsilon^{-2}$ for $m \geqslant 2$. In summary, for $C_1\varepsilon^{-(1+1/m+\delta)} \leqslant t \leqslant C_2\varepsilon^{-2}$ the corrections $\operatorname{Im}\Omega_1^{(2)}$ and $\operatorname{Im}\Omega_2^{(2)}$ play no role for the envelope shifts.

The assertion of Theorem 4.1 obviously follows if we prove that the approximation equations (3.10)–(3.15) possess order $\mathcal{O}(1)$-bounded solutions on the $\mathcal{O}(1/\varepsilon^2)$-time scale. We have to solve three different kinds of equations. The first set of equations, (3.10) and (3.12), describes internal dynamics. Since these two equations are independent of the small parameter $0 < \varepsilon \ll 1$ we have

Lemma 4.2. *Under the assumptions of Theorem 4.1 there exists a time $T_0 > 0$ such that (3.10) has a unique solution*

$$A_1^{(1)} \in C\big([0,T_0], H^{s_A}(m) \cap H^{s_A+m}(0)\big).$$

Proof. We apply the variation of constant formula and use the fact that $i\partial_X^2$ is the generator of a strongly continuous semigroup in $H^{s_A}(m) \cap H^{s_A+m}(0)$, cf. [7]. □

Note that T_0 is independent of the weight. This can be proven like in [12, Lemma 6.4] such that the existence time is determined only by the local existence and uniqueness in H^s-spaces. For completeness we remark that the time $T_0 > 0$ can be made arbitrarily large by using the global well-posedness [15] in the space L^2 which implies the global well-posedness in H^s-spaces for every $s \geqslant 0$.

Since (3.12) is a linearized NLS equation for $A_1^{(2)}$ with $\mathcal{O}(1)$-bounded inhomogeneous terms t_{41} with exactly the same arguments as above we find

Lemma 4.3. *Under the assumptions of Theorem 4.1 the following holds. Let $A_1^{(1)} \in C([0,T_0], H^{s_A}(m) \cap H^{s_A+m}(0))$ be a solution of (3.10). Then for all initial conditions $A_1^{(2)}|_{T=0} \in H^{s_A}(m) \cap H^{s_A+m}(0)$ there exists a unique solution of (3.12) with*

$$A_1^{(2)} \in C\big([0,T_0], H^{s_A-3}(m) \cap H^{s_A-3+m}(0)\big).$$

The loss of regularity for $A_1^{(2)}, A_2^{(2)}$ comes from the inhomogeneous term t_{41} in (3.12). The second group of equations, namely (3.11), (3.13), and (3.14), describes the essential interaction dynamics. By pure integration we find

Lemma 4.4. *Under the assumptions of Theorem 4.1 the following holds. Let $A_1^{(1)}, B_2^{(1)} \in C([0,T_0], H^{s_A}(m) \cap H^{s_A+m}(0))$ be a solution of (3.10). Then*

$$\partial_1\Omega_1^{(1)}, \partial_1\Omega_2^{(1)}, \partial_{X_2}\psi_A, \partial_{X_1}\psi_B, \operatorname{Im}\Omega_1^{(2)}, \operatorname{Im}\Omega_2^{(2)} \in C\big([0,T_0], H^{s_A}(m) \cap H^{s_A+m}(0)\big),$$

and $\Omega_1^{(1)}, \Omega_2^{(1)}, \operatorname{Re}\Omega_1^{(2)}, \operatorname{Re}\Omega_2^{(2)}, \Psi_1^{(1)}, \Psi_2^{(1)} \in C([0,T_0], C_b^{s_A+m})$.

In terms of local existence and uniqueness and $\mathcal{O}(1)$-boundedness of solutions the only nontrivial equation is (3.15) which is a linearized NLS equation for $A_j^{(3)}$ with $\mathcal{O}(1)$-bounded inhomogeneous terms and terms $\varepsilon^{-1}(t_{52})$. Since the last terms are only $\mathcal{O}(\varepsilon^{-1})$ on an $\mathcal{O}(\varepsilon)$-scale w.r.t. T we find

Lemma 4.5. *Assume the conditions of Theorem 4.1. Then there exists a $C > 0$ such that for all $\varepsilon \in (0,1]$ the following holds. System (3.15) with zero initial data has a unique solution $A_1^{(3)}, A_2^{(3)} \in C([0,T_0], H^{s_A-6}(m) \cap H^{s_A-6+m}(0))$. It satisfies*

$$\sup_{0\leqslant T\leqslant T_0} \left\| \left(A_1^{(3)}, A_2^{(3)}\right)(T) \right\|_{H^{s_A-6}(m)\cap H^{s_A-6+m}(0)} \leqslant C.$$

Proof. See [4, Lemma 4.2]. □

Due to the term $\partial_1^3 A_1^{(1)}$ in the equation for $A_1^{(2)}$ we have a loss of regularity from $A_1^{(1)}$ to $A_1^{(2)}$ of three derivatives. Similarly, the loss of regularity of another three derivatives for $A_1^{(3)}$ comes from the term t_{51}. Finally, the terms with the highest derivative in the residual are $\partial_2^2 A_1^{(3)}$ and $\partial_2^2 A_2^{(3)}$ which gives another loss of four spatial derivatives via the right hand sides of the NLS equations. Consequently, we have to choose $s_A - s \geqslant 10$.

5. Estimates for the residual in the original system

Our approximation V^{an} defined in (3.1)–(3.6) has been constructed such that the residual

$$\mathrm{Res}_n(V)(x,t) = -\partial_t V_n(x,t) + D_n(\partial_x) V_n(x,t) + N_n\big((V_j)_{j\in\mathbb{N}}\big)(x,t),$$

is formally of $\mathcal{O}(\varepsilon^6)$. It is the goal of this section to construct an approximation u^{an} out of V^{an} such that the residual

$$\mathrm{Res}\big(u^{an}\big) = -\partial_t^2 u^{an} + \partial_x^2 u^{an} - a u^{an} + b\big(u^{an}\big)^3, \tag{5.1}$$

of the original spatially periodic nonlinear wave equation (5.1) is at least $\mathcal{O}(\varepsilon^{11/2})$ in the Sobolev norm $\|\cdot\|_{H^s}$.

In order to do so we have to undo the transformations in Sections 2.1 and 2.3.

1. In Section 2.1 we started with the Bloch transform $\mathcal{B}$ and defined $\breve{u} = \mathcal{B}u$.
2. In the next step we expanded $\breve{u}$ in terms of the eigenfunctions for fixed l. The mapping to the coefficient functions has been called $\mathcal{P}$ and we defined $\tilde{u} = \mathcal{P}\breve{u}$.
3. In the next step there has been an extension operator $\mathcal{E}_1$ by writing the second order system for $\tilde{u}$ as a first order system and so we defined $\tilde{Z} = \mathcal{E}_1\tilde{u}$. The associated operator we are interested in, is the restriction operator $\mathcal{R}_1$ which takes out every second component out of $\tilde{Z}$ such that $\mathcal{R}_1\mathcal{E}_1 = I$.
4. Finally in Section 2.1 we diagonalized the first order system for $\tilde{Z}$ with help of the operator $\mathcal{Q}$ and so we defined $\tilde{V} = \mathcal{Q}^*\tilde{Z}$.
5. In the first step of Section 2.3 we applied a smooth cut-off operator $\mathcal{C}$ to the Bloch transform $\tilde{V}$. For the interpretation of $\hat{V} = \mathcal{C}\tilde{V}$ as Fourier transform it is then extended by zero to the real axis with an extension operator $\mathcal{E}_2$. Associated to this extension operator is a restriction operator $\mathcal{R}_2$ again defined with the help of a smooth cut-off function.
6. Finally the inverse Fourier transform $\mathcal{F}^{-1}$ has been applied such that $V = \mathcal{F}^{-1}\hat{V}$.

Fix $s \geqslant 1$. Then the analytic properties of these mappings are as follows.

1. The Bloch transform $\mathcal{B}$ is an isomorphism between $H^s(\mathbb{R},\mathbb{C})$ and $L^2([-1/2,1/2),H^s((0,2\pi],\mathbb{C}))$.
2. The mapping $\mathcal{P}$ is an isomorphism between $L^2([-1/2,1/2),H^s((0,2\pi],\mathbb{C}))$ and $L^2([-1/2,1/2),\ell^2(s))$.
3. The extension operator $\mathcal{E}_1$ is a continuous mapping from $L^2([-1/2,1/2),\ell^2(s))$ into $L^2([-1/2,1/2),\ell^2(s)\times\ell^2(s))$ which is identified with $L^2([-1/2,1/2),\ell^2(s))$. The restriction operator $\mathcal{R}_1$ is a continuous mapping vice versa.
4. The diagonalization operator $\mathcal{Q}$ is an isomorphism between $L^2([-1/2,1/2),\ell^2(s))$ and itself.
5. The composition of the extension operator $\mathcal{E}_2$ and cut-off operator $\mathcal{C}$ is a continuous mapping from $L^2([-1/2,1/2),\ell^2(s))$ to $L^2(m)(\mathbb{R},\ell^2(s))$ for every $m\geqslant 0$. The restriction $\mathcal{R}_2$ is continuous vice versa.
6. Finally the inverse Fourier transform $\mathcal{F}^{-1}$ is an isomorphism between $L^2(m)(\mathbb{R},\ell^2(s))$ and $H^m(\mathbb{R},\ell^2(s))$ for every $m\geqslant 0$.

Therefore, we define the approximation u^{an} through

$$u^{an}=\mathcal{B}^{-1}\mathcal{P}^{-1}\mathcal{R}_1\mathcal{Q}\mathcal{R}_2\mathcal{F}V^{an}. \tag{5.2}$$

As a consequence of the previous properties of the involved mappings and Lemma 2.6 it follows that

Lemma 5.1. *Let $s\geqslant 2$, $\tilde{m}\geqslant 2$, $s_A\geqslant s+10$, $l_1\neq l_2$, $l_1,l_2>0$, and let $A_j^{(1)}|_{T=0}\in H^{s_A}(\tilde{m})\cap H^{s_A+\tilde{m}}(0)$ for $j=1,2$. Then for all $T_0>0$ there exist $\varepsilon_0>0$, $C>0$ such that for all $\varepsilon\in(0,\varepsilon_0)$ we have*

$$\sup_{t\in[0,T_0/\varepsilon^2]}\left\|\operatorname{Res}\left(u^{an}\right)\right\|_{H^s}\leqslant C\varepsilon^{11/2}.$$

Since the Fourier and Bloch transforms are also continuous from L^1 to C_b^0 it also follows that $V^{an}\in H^m(\mathbb{R},\ell^2(s))$ of order $\mathcal{O}(\varepsilon)$ implies that $u^{an}\in C_b^{m-1}(\mathbb{R},\mathbb{C})$ is of order $\mathcal{O}(\varepsilon)$, too. This property is needed in Section 6 in order to ensure that the constant C_{an} defined there is of order $\mathcal{O}(1)$.

6. Approximation result

We are now ready to prove the following approximation theorem.

Theorem 2. *Let $A_j^{(1)}\in C([0,T_0],H^{s_B}(2))$ for $j=1,2$ be solutions of the NLS equation (3.10) and let u^{an} be the formal approximation defined in (5.2). Then there exist $\varepsilon_0>0$ and $C>0$ such that for all $\varepsilon\in(0,\varepsilon_0)$ we have solutions u of (1.1) with*

$$\sup_{t\in[0,T_0/\varepsilon^2]}\left\|u-u^{an}\right\|_{C_b^{s-1}}\leqslant C\varepsilon^{7/2}.$$

Proof. The solution u of the spatially periodic Klein–Gordon equation (1.1) is a sum of the approximation $u^{an}=\varepsilon U^{an}$ and the error $\varepsilon^\delta R$, i.e. $u=\varepsilon U^{an}+\varepsilon^\delta R$. Plugging this into (1.1) gives the error equation

$$\partial_t^2R=\partial_x^2R-aR+f$$

with

$$f=3\varepsilon^2b\left(U^{an}\right)^2R+3\varepsilon^{\delta+1}bU^{an}R^2+\varepsilon^{2\delta}bR^3+\varepsilon^{-\delta}\operatorname{Res}\left(\varepsilon U^{an}\right).$$

In order to estimate the error R we introduce the energy

$$E(R) = \int_{\mathbb{R}} aR^2 + (\partial_x R)^2 + (\partial_t R)^2 \, dx.$$

According to our assumption $\inf_{x\in\mathbb{R}} a(x) = a_{\inf} > 0$ we have that $\sqrt{E(\cdot)}$ bounds the H^1-norm from above. We find

$$\begin{aligned}\frac{d}{dt}E(R) &= 2\int aR(\partial_t R) + (\partial_x R)(\partial_t \partial_x R) + (\partial_t R)\big(\partial_t^2 R\big)\, dx\\ &= 2\int aR(\partial_t R) - \big(\partial_x^2 R\big)(\partial_t R) + (\partial_t R)\big(\partial_x^2 R - aR + f\big)\, dx\\ &= 2\int (\partial_t R) f \, dx \leqslant 2\|\partial_t R\|_{L^2}\|f\|_{L^2} \leqslant 2E(R)^{1/2}\|f\|_{L^2}.\end{aligned}$$

We can estimate f as

$$\|f\|_{H^1} \leqslant 3\varepsilon^2 c_2 C_{an}^2 \|R\|_{H^1} + 3\varepsilon^{\delta+1} c_2 C_{an}\|R\|_{H^1}^2 + \varepsilon^{2\delta}\|R\|_{H^1}^3 + C_{res}\varepsilon^{res-\delta},$$

where $C_{an} := \|U^{an}\|_{C_b^0} = \mathcal{O}(1)$, where $c_2 = \sup_{x\in\mathbb{R}} |b(x)|$ and where $C_{res} > 0$ is defined by $\|\mathrm{Res}(\varepsilon U^{an})\|_{H^1} \leqslant C_{res}\varepsilon^{res}$. Thus we get

$$\begin{aligned}\frac{d}{dt}E(R) &\leqslant 2E(R)^{1/2}\big[3\varepsilon^2 c_2 C_{an}^2\|R\|_{H^1} + 3\varepsilon^{\delta+1}c_2 C_{an}\|R\|_{H^1}^2 + \varepsilon^{2\delta}c_2\|R\|_{H^1}^3 + C_{res}\varepsilon^{res-\delta}\big]\\ &\leqslant 2E(R)^{1/2}\big[3\varepsilon^2 c_2 C_{an}^2 E(R)^{1/2} + 3\varepsilon^{\delta+1}c_2 C_{an}E(R) + \varepsilon^{2\delta}c_2 E(R)^{3/2} + C_{res}\varepsilon^{res-\delta}\big]\\ &= 2\big[3\varepsilon^2 c_2 C_{an}^2 E(R) + 3\varepsilon^{\delta+1}c_2 C_{an}E(R)^{3/2} + \varepsilon^{2\delta}c_2 E(R)^2 + C_{res}\varepsilon^{res-\delta}E(R)^{1/2}\big].\end{aligned}$$

We introduce the rescaled time $T = \varepsilon^2 t$. Using the inequality $|y| \leqslant 1 + y^2$ finally gives the estimate

$$\begin{aligned}\frac{d}{dT}E(R) \leqslant\ & C_{res}\varepsilon^{res-\delta-2} + \big(3c_2 C_{an}^2 + 3\varepsilon^{\delta-1}c_2 C_{an} + C_{res}\varepsilon^{res-\delta-2}\big)E(R)\\ &+ \big(3\varepsilon^{\delta-1}c_2 C_{an} + \varepsilon^{2\delta-2}\big)E(R)^2\end{aligned}$$

Since $res = 11/2$ and $\delta = 7/2$ this inequality is of the form

$$\frac{d}{dT}E(R) \leqslant \alpha(\varepsilon) + \beta(\varepsilon)E(R) + \gamma(\varepsilon)\varepsilon^{5/2}E(R)^2$$

with $\alpha(\varepsilon), \beta(\varepsilon), \gamma(\varepsilon) = \mathcal{O}(1)$ for $\varepsilon \to 0$. Hence a simple application of Gronwall's inequality gives the $\mathcal{O}(1)$-boundedness of $E(R)$ for all $T \in [0, T_0]$ and all $\varepsilon \in (0, \varepsilon_0)$ if $\varepsilon_0 > 0$ is chosen sufficiently small. Hence we bounded the H^1-norm of the error function R with an $\mathcal{O}(1)$-bound for all $t \in [0, T_0/\varepsilon^2]$. In order to bound the H^s-norm we use the energy

$$E(R) = \sum_{j=0}^{s-1}\int_{\mathbb{R}} a\big(\partial_t^j R\big)^2 + \big(\partial_t^j \partial_x R\big)^2 + \big(\partial_t^{j+1} R\big)^2\, dx.$$

By taking the derivatives w.r.t. t the x-dependent coefficients can be handled as above. The estimates for the derivatives w.r.t. x can be obtained with the error equations which we write as

$$\partial_x^2 R = \partial_t^2 R + aR - f.$$

Finally, the application of Sobolev's embedding theorem concludes the proof. □

Remark 6.1. The proof of Theorem 2 is the same as for the description of a single wave packet by the NLS approximation. Hence the restriction to a wave equation with cubic nonlinearity in this paper is only for clarity. With the same procedure wave equations with quadratic nonlinearities can be handled. The associated approximation theorem in the quadratic case with spatially periodic coefficients can be found in [3].

7. Summary and transfer of the result to physical space

In the last section we have shown that u^{an} given by (5.2) provides a good approximation of the solutions u of the spatially periodic nonlinear wave equation (1.1) over the natural time scale. However, in contrast to V^{an}, for which internal and interaction dynamics of the pulses are well separated, for u^{an} there is a mixture of internal and interaction dynamics due to transformations involved. Hence, it is the purpose of this section to write down an expansion of u^{an} and to give an interpretation of the formulas w.r.t. to possible measurements of the phase and envelope shift in possible experiments.

As already explained it is not possible to write down the inverse Fourier transform of the ansatz V^{an} explicitly. However, for our purposes this is not necessary. It is sufficient to compute the inverse Fourier transform and the other transformations involved of an expansion of the ansatz V^{an} w.r.t. the small perturbation parameter $0 < \varepsilon \ll 1$.

Expanding V^{an} w.r.t. ε gives

$$\begin{aligned} v_1^{an}(x,t) = \sum_{j=1,2} E_j\big(&\varepsilon A_j^{(1)}(X_j,T) + \varepsilon^2 A_j^{(2)}(X_j,T) + \varepsilon^2 A_j^{(1)}(X_j,T) i\Omega_j^{(1)}(X_{3-j},T) \\ &+ \varepsilon^3 A_j^{(3)}(X_j,T) + \varepsilon^3 A_j^{(2)}(X_j,T) i\Omega_j^{(1)}(X_{3-j},T) \\ &+ \varepsilon^3 A_j^{(1)}(X_j,T) i\Omega_j^{(2)}(X_{3-j},T) + \varepsilon^3 \partial_1 A_j^{(1)}(X_j,T)\Psi_j^{(1)}(X_{3-j},T) \\ &+ \mathcal{O}\big(\varepsilon^4\big)\big) + \varepsilon^3 M^1_{mixed,int} + \varepsilon^3 M^1_{mixed,inter}, \end{aligned} \tag{7.1}$$

where $E_j = e^{il_j x + i\omega_1(l_j)t}$ and $X_j = \varepsilon(x - c_j t)$. We find a similar expression for $v_{-1}^{an}(x,t)$. Moreover, we have

$$v_j^{an}(x,t) = \varepsilon^3 M^j_{mixed,int} + \varepsilon^3 M^j_{mixed,inter}.$$

We have split the mixed terms into mixed terms $M^1_{mixed,int}$ describing the internal dynamics and into terms $M^1_{mixed,inter}$ describing interaction dynamics. These last terms vanish rapidly to zero outside the domain of interaction.

We have

$$u^{an}(x,t) = \int_{-1/2}^{1/2} \big(\chi(l)\hat{v}_1(l,t) f_1(l,x) + \chi(l)\hat{v}_{-1}(l,t) f_1(l,x) + \mathcal{O}\big(\varepsilon^3\big)\big) e^{ilx}\, dl.$$

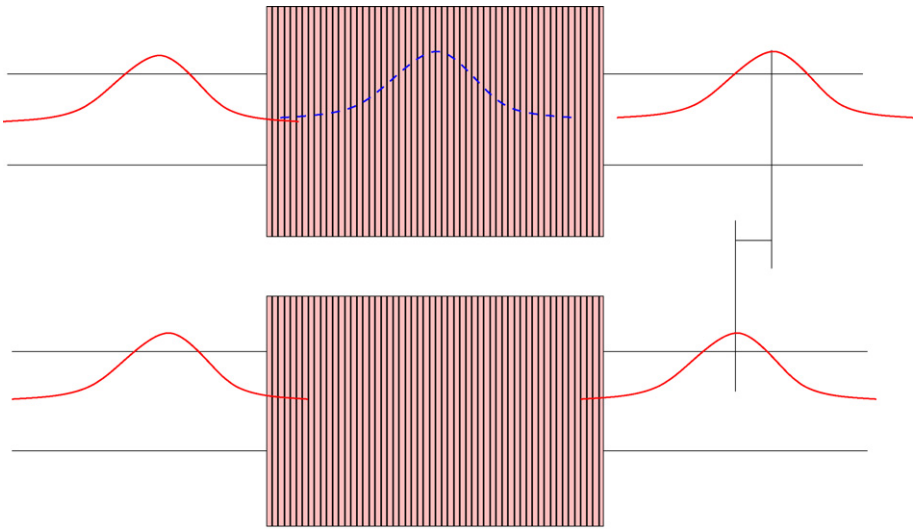

Fig. 4. Separating internal and interaction dynamics by comparing the interaction experiment with a non-interaction experiment.

Using the Fourier transform of (7.1) and

$$f_1(l,x) = f_1(l_j,x) + i(l-l_j)f_1'(l_j,x) - (l-l_j)^2 f_1''(l_j,x)/2 + \mathcal{O}\big((l-l_j)^3\big)$$

where $' = \partial_l$ shows that

$$\begin{aligned} u^{an}(x,t) = \sum_{j=1,2} E_j\big(&\varepsilon A_j^{(1)}(X_j,T) f_1(l_j,x) + \varepsilon^2 A_j^{(2)}(X_j,T) f_1(l_j,x) \\ &+ \varepsilon^2 A_j^{(1)}(X_j,T) i\Omega_j^{(1)}(X_{3-j},T) f_1(l_j,x) \\ &+ \varepsilon^2 \partial_1 A_j^{(1)}(X_j,T) f_1'(l_j,x) + \varepsilon^3 A_j^{(3)}(X_j,T) f_1(l_j,x) \\ &+ \varepsilon^3 A_j^{(2)}(X_j,T) i\Omega_j^{(1)}(X_{3-j},T) f_1(l_j,x) \\ &+ \varepsilon^3 A_j^{(1)}(X_j,T) i\Omega_j^{(2)}(X_{3-j},T) f_1(l_j,x) \\ &+ \varepsilon^3 \partial_1 A_j^{(1)}(X_j,T) \Psi_j^{(1)}(X_{3-j},T) f_1(l_j,x) + \varepsilon^3 \partial_1^2 A_j^{(1)}(X_j,T) f_1''(l_j,x)/2 \\ &+ \varepsilon^3 \partial_1 A_j^{(2)}(X_j,T) f_1'(l_j,x) + \varepsilon^3 \partial_1 A_j^{(1)}(X_j,T) i\Omega_j^{(1)}(X_{3-j},T) f_1'(l_j,x) \\ &+ \varepsilon^3 A_j^{(1)}(X_j,T) i\partial_1 \Omega_j^{(1)}(X_{3-j},T) f_1'(l_j,x) \\ &+ \mathcal{O}\big(\varepsilon^4\big)\big) + \varepsilon^3 M^1_{mixed,int} + \varepsilon^3 M^1_{mixed,inter} + c.c., \end{aligned}$$

using the fact that there is only a small difference between Bloch and Fourier transform for functions concentrated at various wave numbers, cf. Lemma 2.7.

At a first sight this expression looks rather useless. However in a possible experiment for a detection of a standing pulse in a photonic crystal we measure only the first pulse. The internal dynamics can be separated by the second experiment with the same device but without the standing pulse. See Fig. 4. Moreover, we can neglect all terms which are only spatially localized and vanish after the interaction, like $\Omega_j^{(2)}$ and $M^1_{mixed,inter}$. Hence the remaining interaction terms are given by

$$\begin{aligned}
&\sum_{j=1,2} E_j\big(\varepsilon^2 A_j^{(1)}(X_j,T)\Omega_j^{(1)}(X_{3-j},T)f_1(l_j,x)\\
&\quad+\varepsilon^3 A_j^{(2)}(X_j,T)\Omega_j^{(1)}(X_{3-j},T)f_1(l_j,x)\\
&\quad+\varepsilon^3\partial_1 A_j^{(1)}(X_j,T)\Psi_j^{(1)}(X_{3-j},T)f_1(l_j,x)\\
&\quad+\varepsilon^3\partial_1 A_j^{(1)}(X_j,T)\Omega_j^{(1)}(X_{3-j},T)f_1'(l_j,x)\\
&\quad+\varepsilon^3 A_j^{(1)}(X_j,T)\partial_1\Omega_j^{(1)}(X_{3-j},T)f_1'(l_j,x)\\
&\quad+\mathcal{O}\big(\varepsilon^4\big)\big)+c.c. \qquad (7.2)
\end{aligned}$$

For the spatially homogeneous Klein–Gordon equation (1.2) we have $f_1(l_j,x)=1$ and $f_1'(l_j,x)=0$. In this case the procedure to measure the overall carrier shift and the overall envelope position shift has been described in detail in [5]. In the homogeneous case the phase shift can easily be determined by comparing the zeroes of the interacting and of the non-interacting pulse of the second experiment. In non-homogeneous case the Bloch function $f_1(l_j,x)$ has to be taken into account. However, in general their zeroes are not equidistant which makes the experimental measurement of the phase shift a hard task.

On the other hand the manipulation of a standing pulse by an interaction mechanism can be designed by using the equations that we derived for the main interaction effects, namely, the formula for the **carrier shift**

$$\tilde{\Omega}_1^{(1)}(X_2,T)=\frac{6}{\omega_1'(k_1)-\omega_1'(k_2)}\left(\int_0^{2\pi} b(x)\big|f_1(l_1,x)\big|^2\big|f_1(l_2,x)\big|^2\,dx\right)\int_{-\infty}^{X_2}\big|A_2^{(1)}(\xi,T)\big|^2\,d\xi$$

and the formula for the **envelope position shift** which adds in (7.2) to the next order manipulation.

Acknowledgments

This research was partially supported by the Graduiertenkolleg 1294: Analysis, simulation, and design of nanotechnological processes, and the Karlsruhe House of Young Scientists (KHYS) at the Karlsruhe Institute of Technology (KIT). The authors are grateful for discussions with Kurt Busch, Christopher Chong, and Hannes Uecker. Martina Chirilus-Bruckner is grateful for the stimulating atmosphere at the Centrum Wiskunde & Informatica in Amsterdam where a part of this article has been completed.

References

[1] A. Babin, A. Figotin, Linear superposition in nonlinear wave dynamics, Rev. Math. Phys. 18 (9) (2006) 971–1053.
[2] C. Blank, M. Chirilus-Bruckner, V. Lescarret, G. Schneider, Breather solutions in periodic media, Comm. Math. Phys. 302 (2011) 815–841.
[3] K. Busch, G. Schneider, L. Tkeshelashvili, H. Uecker, Justification of the nonlinear Schrödinger equation in spatially periodic media, Z. Angew. Math. Phys. 57 (6) (2006) 905–939.
[4] M. Chirilus-Bruckner, G. Schneider, H. Uecker, On the interaction of NLS-described modulating pulses with different carrier waves, Math. Methods Appl. Sci. 30 (15) (2007) 1965–1978.
[5] M. Chirilus-Bruckner, C. Chong, G. Schneider, H. Uecker, Separation of internal and interaction dynamics for NLS-described wave packets with different carrier waves, J. Math. Anal. Appl. 347 (1) (2008) 304–314.
[6] M. Chirilus-Bruckner, Nonlinear interaction of pulses, Dissertation, Karlsruhe, Univ., 2009, URL: http://digbib.ubka.uni-karlsruhe.de/volltexte/1000012894.
[7] W. Craig, T. Kappeler, W. Strauss, Microlocal dispersive smoothing for the Schrödinger equation, Comm. Pure Appl. Math. 48 (1995) 769–860.
[8] P. Kirrmann, G. Schneider, A. Mielke, The validity of modulation equations for extended systems with cubic nonlinearities, Proc. Roy. Soc. Edinburgh Sect. A 122 (1–2) (1992) 85–91.

[9] V. Lescarret, C. Blank, M. Chirilus-Bruckner, C. Chong, G. Schneider, Standing generalized modulating pulse solutions for a nonlinear wave equation in periodic media, Nonlinearity 22 (8) (2009) 1869–1898.
[10] M. Oikawa, N. Yajima, A perturbation approach to nonlinear systems. II. Interaction of nonlinear modulated waves, J. Phys. Soc. Japan 37 (2) (1974) 486–496.
[11] R.D. Pierce, C.E. Wayne, On the validity of mean-field amplitude equations for counterpropagating wavetrains, Nonlinearity 8 (5) (1995) 769–779.
[12] G. Schneider, C.E. Wayne, The long-wave limit for the water wave problem. I. The case of zero surface tension, Comm. Pure Appl. Math. 53 (2000) 1475–1535.
[13] G. Schneider, H. Uecker, M. Wand, Interaction of modulated pulses in nonlinear oscillator chains, J. Difference Equ. Appl. 17 (3) (2011) 279–298.
[14] L. Tkeshelashvili, S. Pereira, K. Busch, General theory of nonresonant wave interaction: Giant soliton shift in photonic band gap materials, Europhys. Lett. 68 (2004) 205.
[15] Y. Tsutsumi, L^2-solutions for nonlinear Schrödinger equations and nonlinear groups, Funkcial. Ekvac. 30 (1987) 115–125.
[16] M. Reed, B. Simon, Methods of Modern Mathematical Physics. IV: Analysis of Operators, Academic Press Inc., 1978.